

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 05:06:07 ; Search time 206.5 Seconds  
(without alignments)  
261.447 Million cell updates/sec

Title: US-09-927-585a-2

Perfect score: 20  
Sequence: 1 tgcattcgcgtctactctg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	14	AAQ48541	HPV E6/7 region pr
2	20	100.0	20	14	AAQ48565	HPV E6/7 region pr
3	20	100.0	20	14	AAQ48565	Low hazard strain
4	20	100.0	731	20	AAQ91601	Partial nucleotide
5	20	100.0	1107	18	AAQ59430	Human papillomavir
6	20	100.0	1128	18	AAQ59431	Human papillomavir
7	20	100.0	1398	18	AAQ59432	Human papillomavir
8	20	100.0	4770	18	AAQ59450	Human papillomavir

9	20	100.0	8010	17	AAQ13413	Human papillomavir
10	18	90.0	23	25	ABX12328	Fluorescent multipl
11	15.8	79.0	281	22	AAQ29985	Human lung antigen
12	15.8	79.0	281	22	AAQ29985	Genomic sequence #
13	15.8	79.0	287	22	AAQ29986	Human lung antigen
14	15.8	79.0	287	22	AAQ29986	Novel lunc antigen
15	15.8	79.0	287	25	ABX28028	Aspergillus oryzae
16	15.8	79.0	4682	25	ABX28028	Aspergillus oryzae
17	15.4	77.0	640	23	ABX45803	Human prostate exp
18	15.4	77.0	650	23	ABX45803	Human prostate exp
19	15.4	77.0	1111	23	AAQ16004	DNA encoding novel
20	15.4	77.0	1368	24	ABX68579	Streptococcus poly
21	15.4	77.0	1368	24	ABX68579	Streptococcus poly
22	15.4	77.0	2564	23	ABX30275	Human prostate exp
23	15.4	77.0	3340	21	AAQ79267	Human prostate exp
24	15.4	77.0	213561	24	ABX71527	Pinus radiata cell
25	15.2	76.0	20	14	AAQ48545	Streptococcus poly
26	15.2	76.0	245	25	ABX63752	HPV E6/7 region pr
27	15.2	76.0	720	10	AAQ91603	Human CDNA #752 d1
28	15.2	76.0	767	24	AAQ48310	Partial nucleotide
29	15.2	76.0	1715	22	ABX18187	HPV11-E6/E7 region
30	15.2	76.0	1716	22	ABX18187	Human nervous syst
31	15.2	76.0	12587	22	AAQ87087	Human nervous syst
32	15.2	76.0	65921	21	AAQ289046	Human immune/haema
33	15.2	76.0	349980	22	AAQ86431	Human nibrin DNA.
34	14.8	74.0	94	21	AAQ23951	Pyrococcus abyssi
35	14.8	74.0	1285	19	AAQ59580	Human secreted pro
36	14.8	74.0	1285	19	AAQ59580	Human secreted pro
37	14.8	74.0	1299	23	ABX18657	Human CDNA #1 for
38	14.8	74.0	2294	23	ABX18657	Drosophila melanog
39	14.8	74.0	2537	22	AAQ125274	Drosophila melanog
40	14.8	74.0	3330	23	ABX17687	Human CDNA sequenc
41	14.8	74.0	3363	23	ABX18656	Drosophila melanog
42	14.8	74.0	4165	23	ABX05520	Drosophila melanog
43	14.8	74.0	4520	23	ABX05532	Drosophila melanog
44	14.8	74.0	4554	23	ABX10638	Drosophila melanog
45	14.8	74.0	5626	23	ABX21562	Drosophila melanog

# ALIGNMENTS

AAQ48541	standard; DNA; 20 BP.
AAQ48541	
22-FEB-1994	(first entry)
HPV E6/7 region probe.	
Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.	
Synthetic.	
JP05192200-A.	
03-AUG-1993.	
19-AUG-1991.	91JP-0230839.
20-AUG-1990.	90JP-0217067.
(TAKI ) TAKARA SHUZO CO LTD.	
WPI; 1993-277497/35.	
detecting benign and/or malignant human papilloma virus - by	
detecting DNA sequence of E6 and/or E7 region of human papilloma	
virus	
Disclosure; Page 13; 18pp; Japanese.	

XX The probe is used to detect benign and/or malignant human papilloma  
CC virus. The probe binds to the E6 and/or E7 region of the virus.  
XX  
SQ Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTCTACCTG 20  
DB 1 TGCTAATTCGGTGTCTACCTG 20

RESULT 2  
AAQ48565  
ID AAQ48565 standard; DNA; 20 BP.  
XX  
AC AAQ48565;  
XX  
DT 22-FEB-1994 (first entry)  
XX  
DE HPV E6/7 region probe.  
XX  
KM Human papilloma virus; HPV; E6; E7; benign; malignant; probe: ss.  
XX  
OS Synthetic.  
XX  
PN JP05192200-A.  
XX  
PD 03-AUG-1993.  
XX  
PF 19-AUG-1991; 91JP-0230839.  
XX  
PR 20-AUG-1990; 90JP-0217067.  
XX  
PA (TAKI ) TAKARA SHUZO CO LTD.  
XX  
DR WPI; 1993-277497/35.  
XX  
PT Detecting benign and/or malignant human papilloma virus - by  
PT detecting DNA sequence of E6 and/or E7 region of human papilloma  
PT virus  
XX  
PS Disclosure; Page 15; 18pp; Japanese.  
XX  
CC The probe is used to detect benign and/or malignant human papilloma  
CC virus. The probe binds to the E6 and/or E7 region of the virus.  
XX  
SQ Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTCTACCTG 20  
DB 1 TGCTAATTCGGTGTCTACCTG 20

RESULT 3  
AAS18613  
ID AAS18613 standard; DNA; 20 BP.  
XX  
AC AAS18613;  
XX  
DT 26-FEB-2002 (first entry)  
XX  
DE Low hazard strain human papillomavirus (HPV) E6, PCR primer.  
XX  
KM Human papillomavirus; HPV; early gene; E6; PCR primer;  
KM low hazard strain; ss.

XX Human papillomavirus.  
OS  
XX  
XX  
PN WO200185994-A2.  
XX  
XX  
PD 15-NOV-2001.  
XX  
XX  
PF 02-MAY-2001; 2001WO-IB00771.  
XX  
XX  
PR 05-MAY-2000; 2000IT-MO00091.  
XX  
XX  
PA (BIOA-) BIOANALISI CENT SUD DI PERSEU SINIBLADO.  
XX  
PI Perseu S, De Montis A, Floris MM;  
XX  
DR WPI; 2002-075253/10.  
XX  
XX  
PT Identifying viral DNA of human papilloma virus in cellular material  
PT collected from an individual, by using polymerase and identifying  
PT sequences complementary to regions E of the viral genome using a  
PT reagent  
XX  
PS Claim 11; Page 12; 17pp; English.  
XX  
XX  
CC The invention describes a novel method of identifying viral DNA of human  
CC papillomavirus (HPV) in cellular material collected from an individual.  
CC The method comprises extracting DNA from the cellular material,  
CC amplifying the DNA by using a polymerase, and identifying sequences  
CC complementary to sequences in the regions E (early genes) e.g. E6 and E7  
CC of the HPV genome using an identifying reagent. The presence of HPV in  
CC cells is now based on identification of a nucleotide sequence in the L1  
CC region of the genome. Identification of E6, E7 and L1 make it possible to  
CC identify the presence of HPV virus in every stage of infection and to  
CC distinguish between virus strains having low hazard and virus strains  
CC having middle/high hazard. This is the primer used to identify the E6  
CC region of the low hazard HPV strains in infected cells, described in the  
CC method of the invention.  
XX  
SQ Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTCTACCTG 20  
DB 1 TGCTAATTCGGTGTCTACCTG 20

RESULT 4  
AAN91601  
ID AAN91601 standard; DNA; 731 BP.  
XX  
XX  
AC AAN91601;  
XX  
XX  
DT 17-JUL-1990 (first entry)  
XX  
XX  
DE Partial nucleotide sequence (5' end) of human papilloma virus (HPV)  
DE type 6 (HPV-6).  
XX  
XX  
KM Human papilloma virus; type 6; in situ hybridisation assay;  
KM cellular smear; benign cervical wart; cervical cancer.  
XX  
OS Human papilloma virus.  
XX  
PN WO8902934-A.  
XX  
XX  
PD 06-APR-1989.  
XX  
XX  
PF 30-SEP-1988; 88WO-US03367.  
XX  
XX  
PR 02-OCT-1987; 87US-0103979.  
XX

PA (MIR-) MICROPROBE CORP.  
XX  
PI Schwartz DE, Adams TH;  
XX  
DR WPI: 1989-114406/15.  
XX  
PT Hybridisation test for human papilloma virus in cell smears -  
PT by reaction with long labelled probe specific for particular  
PT virus types, esp. for examining cervical smears  
XX  
PS Disclosure: 39pp; English.  
XX  
CC The patent is for a rapid in situ hybridisation assay for detecting and  
CC typing human papilloma virus (HPV) in non-frozen cellular smears fixed to  
CC a support in absence of aldehyde-based crosslinking reagents. The assay  
CC comprises: (1) combining nucleic acid in the sample with at least one  
CC detectable probe able to hybridise with 1 or more HPV types; and (2)  
CC detecting presence or absence of hybrid complexes. Opt. several probes  
CC are used, eg one for HPV types 6 and 11, associated with benign warts,  
CC and one for types 16, 18, 31, 33 and 35, associated with cervical  
CC cancer. The assay can differentiate between HPV types. It is esp. used  
CC as a secondary test. The probes can be synthesised or cloned.  
XX  
SQ Sequence 731 BP; 245 A; 147 C; 159 G; 180 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 10; Length 731;  
Best Local Similarity 100.0%; Pred. No. 0.44;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 TGCTAATTCGCTCTACTG 20  
DB 400 TGCTAATTCGCTCTACTG 419  
  
RESULT 5  
AAT59430  
ID AAT59430 standard; DNA: 1107 BP.  
XX  
AC AAT59430;  
XX  
DT 26-MAY-1997 (first entry)  
XX  
DE Human papillomavirus 6b E6/E5a/E4 DNA cassette CSL690.  
XX  
KM HPV; polyprotein; vaccine; wart; condylomata acuminata;  
KM cervical cancer; ss.  
XX  
OS Human papillomavirus type 6b.  
XX  
PN WO9705164-A1.  
XX  
PD 13-FEB-1997.  
XX  
PF 26-JUL-1996; 96WO-AU00473.  
XX  
PR 27-JUL-1995; 95AU-0004439.  
XX  
PS (CSLC-) CSL LTD.  
PA (UYOU) UNIV QUEENSLAND.  
XX  
PI Cox JC, Edwards SJ, Frazer I, Margetts MB, McMillan NAJ;  
PI Moloney MBH, Webb EA, Williams MP;  
XX  
DR WPI: 1997-145619/13.  
DR P-PSDB; AAW01808.  
XX  
PT Papilloma virus early open reading frame polyprotein constructs -  
PT useful for treating, e.g. genital warts and cervical cancer  
XX  
PS Example 1; Page 22-23; 62pp; English.  
XX  
CC DNA construct CSL690 (AAT59430) codes for an E6/E5a/E4 polyprotein  
CC (AAW01808), plus C-terminal hexahistidine tag, of human

CC papillomavirus 6b (HPV6b). It was obtd. by separate PCR  
CC amplification of the E6, E4 and E5a genes (see also AAT59412-17) of  
CC HCv6b and their cloning into vector pSP70. The resulting  
CC E6/E5a/E4 cassette was incorporated into pGEX-STOP to allow prodn.  
CC of the HPV6b polyprotein in E. coli. The E6/E5a/E4 DNA construct,  
CC or the encoded polypeptide, can be used as a vaccine to generate a  
CC humoral and/or cellular immune response against HPV, esp. for  
CC treating genital warts and cervical cancer.  
XX  
SQ Sequence 1107 BP; 337 A; 268 C; 213 G; 289 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 18; Length 1107;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 TGCTAATTCGCTCTACTG 20  
DB 299 TGCTAATTCGCTCTACTG 318  
  
RESULT 6  
AAT59431  
ID AAT59431 standard; DNA: 1128 BP.  
XX  
AC AAT59431;  
XX  
DT 26-MAY-1997 (first entry)  
XX  
DE Human papillomavirus 6b E6/E7/E4 DNA cassette CSL760.  
XX  
KM HPV; polyprotein; vaccine; wart; condylomata acuminata;  
KM cervical cancer; ss.  
XX  
OS Human papillomavirus type 6b.  
XX  
PN WO9705164-A1.  
XX  
PD 13-FEB-1997.  
XX  
PF 26-JUL-1996; 96WO-AU00473.  
XX  
PR 27-JUL-1995; 95AU-0004439.  
XX  
PS (CSLC-) CSL LTD.  
PA (UYOU) UNIV QUEENSLAND.  
XX  
PI Cox JC, Edwards SJ, Frazer I, Margetts MB, McMillan NAJ;  
PI Moloney MBH, Webb EA, Williams MP;  
XX  
DR WPI: 1997-145619/13.  
DR P-PSDB; AAW01809.  
XX  
PT Papilloma virus early open reading frame polyprotein constructs -  
PT useful for treating, e.g. genital warts and cervical cancer  
XX  
PS Example 1; Page 24-25; 62pp; English.  
XX  
CC DNA construct CSL760 (AAT59431) codes for an E6/E7/E4 polyprotein  
CC (AAW01809) of human papillomavirus 6b (HPV6b), plus a C-terminal  
CC hexahistidine tag. It was obtd. by separate PCR amplification of  
CC the E6, E4 and E7 genes (see also AAT59412-15 and AAT59418-19) of  
CC HCv6b and their cloning into vector pSP70. The resulting  
CC E6/E7/E4 cassette was incorporated into pGEX-STOP to allow prodn.  
CC of the HPV6b polyprotein in E. coli. The E6/E7/E4 DNA construct,  
CC or the encoded polypeptide, can be used as a vaccine to generate a  
CC humoral and/or cellular immune response against HPV, esp. for  
CC treating genital warts and cervical cancer.  
XX  
SQ Sequence 1128 BP; 359 A; 274 C; 240 G; 255 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 18; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      1 TGCCTAATTCGGTGTACTCTG 20
XX      |||||||
DB      299 TGCCTAATTCGGTGTACTCTG 318

RESULT 7
AAT59432
ID      AAT59432 standard; DNA; 1398 BP.
XX
AC      AAT59432;
XX
DT      26-MAY-1997 (first entry)
XX
DE      Human papillomavirus 6b E6/E7/E5a/E4 DNA cassette CSL673.
XX
KM      HPV; polypotein; vaccine; wart; condylomata acuminata;
XX      cervical cancer; ss.
XX
OS      Human papillomavirus type 6b.
XX
PN      WO9705164-A1.
XX
PD      13-FEB-1997.
XX
PF      26-JUL-1996; 96WO-AU00473.
XX
PR      27-JUL-1995; 95AU-0004439.
XX
PA      (CSLC-) CSL LTD.
XX      (UYOU ) UNIV QUEENSLAND.
XX
PI      Cox JC, Edwards SJ, Frazer I, Margetts MB, McMillan NAJ;
XX      Moloney MBH, Webb EA, Williams MP;
XX      MPI: 1997-145619/13.
XX      P-PSDB; AAM01810.
XX
DR      Papilloma virus early open reading frame polypotein constructs -
XX      useful for treating, e.g. genital warts and cervical cancer
XX
PS      Example 1; Page 26-28; 62pp; English.
XX
CC      DNA construct CSL673 (AAT59432) codes for an E6/E7/E5a/E4 polypotein
XX      (AAM01810) of human papillomavirus 6b (HPV6b), plus a C-terminal
XX      hexahistidine tag. It was obtd. by separate PCR amplification of
XX      the E6, E4, E5a and E7 genes (see also AAT59412-19 of HCV6b and their
XX      cloning into vector pSP70. The resulting E6/E7/E5a/E4 cassette
XX      was incorporated into pEX-STOP to allow prodn. of the HPV6b
XX      polypotein in E. coli. The E6/E7/E5a/E4 DNA construct, or the
XX      encoded polypeptide, can be used as a vaccine to generate a
XX      humoral and/or cellular immune response against HPV, esp. for
XX      treating genital warts and cervical cancer.
XX
SQ      Sequence 1398 BP; 429 A; 330 C; 283 G; 356 T; 0 other;
XX
Query Match      100.0%; Score 20; DB 18; Length 1398;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 TGCCTAATTCGGTGTACTCTG 20
XX      |||||||
DB      299 TGCCTAATTCGGTGTACTCTG 318

RESULT 8
AAT59450
ID      AAT59450 standard; DNA; 4770 BP.
XX
AC      AAT59450;
XX
DT      31-MAY-1997 (first entry)
XX

```

```

DE      Human papillomavirus 6 polypotein DNA.
XX
KM      HPV; polypotein; vaccine; wart; condylomata acuminata;
XX      cervical cancer; ds.
XX
OS      Human papillomavirus type 6.
XX
FH      Key      Location/Qualifiers
XX      CDS      1..4764
XX      FT      /*tag= a
XX      FT      /product= hexahis-HPV6 polypotein
XX      FT      1..108
XX      FT      /*tag= b
XX      FT      /function= Tag used for protein purification
XX      FT      /note= "hexahis leader sequence form pTCHisA"
XX      FT      115..1218
XX      FT      /*tag= "c
XX      FT      /note= "HPV-6 E2"
XX      FT      1225..1551
XX      FT      /*tag= d
XX      FT      /note= "HPV-6 E4"
XX      FT      1558..1830
XX      FT      /*tag= e
XX      FT      /note= "HPV-6 E5a"
XX      FT      1837..2052
XX      FT      /*tag= f
XX      FT      /note= "HPV-6 E5b"
XX      FT      2059..2508
XX      FT      /*tag= g
XX      FT      /note= "HPV-6 E6"
XX      FT      2515..2808
XX      FT      /*tag= h
XX      FT      /note= "HPV-6 E7"
XX      FT      2815..4764
XX      FT      /*tag= i
XX      FT      /note= "HPV-6 E1"
XX
XX
XX      WO9705164-A1.
XX
XX      13-FEB-1997.
XX
XX
XX      26-JUL-1996; 96WO-AU00473.
XX
XX      27-JUL-1995; 95AU-0004439.
XX
XX      (CSLC-) CSL LTD.
XX      (UYOU ) UNIV QUEENSLAND.
XX
XX      Cox JC, Edwards SJ, Frazer I, Margetts MB, McMillan NAJ;
XX      Moloney MBH, Webb EA, Williams MP;
XX      MPI: 1997-145619/13.
XX      P-PSDB; AAM01816.
XX
XX      Papilloma virus early open reading frame polypotein constructs -
XX      useful for treating, e.g. genital warts and cervical cancer
XX
PS      Example 6; Page 35-40; 62pp; English.
XX
CC      A DNA construct (AAT59450) in vector pTCHisA codes for an
XX      E2/E4/E5a/E5b/E6/E7/E1 polypotein (AAM01810) of human papillomavirus
XX      6 (HPV6), plus an N-terminal hexahistidine tag. It was obtd. by
XX      amplification of the individual early genes (see also AAT59436-49)
XX      and their cloning into pTCHisA. The DNA construct, or the encoded
XX      polypotein, can be used as a vaccine to generate a humoral and/or
XX      cellular immune response against HPV, esp. for treating genital
XX      warts and cervical cancer.
XX
SQ      Sequence 4770 BP; 1597 A; 890 C; 1071 G; 1212 T; 0 other;
XX
Query Match      100.0%; Score 20; DB 18; Length 4770;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY      1  TCGTAATTCGGTGTCTACCTG  20
          |||||
Db     2357 TCGTAATTCGGTGTCTACCTG  2376

```

XX	RESULT 9
XX	AAT13413
XX	ID AAT13413 standard; cDNA; 8010 BP.
AC	AAT13413;
XX	12-JUN-1996 (first entry)
DT	Human papillomavirus type 6a.
XX	DE Human papillomavirus type 6a.
XX	KW Vaccine; immunotherapy; genetic immunisation; condyloma acuminata;
XX	ss.
OS	Human papillomavirus type 6a.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	31..555
FT	/tag= a
FT	/label= ORF-E6
FT	441..827
FT	/tag= b
FT	/label= ORF-E7
FT	716..2782
FT	/tag= c
FT	/label= ORF-E1
FT	2695..3830
FT	/tag= d
FT	/label= ORF-E2
FT	2240..3585
FT	/tag= e
FT	/label= ORF-E4
FT	3888..4163
FT	/tag= f
FT	/label= ORF-E5
FT	4379..5803
FT	/tag= g
FT	/label= ORF-L2
FT	5679..7292
FT	/tag= h
FT	/label= ORF-L1
FT	/product= major capsid protein
XX	
XX	WO9609375-A1.
XX	
XX	28-MAR-1996.
XX	
XX	PD
XX	PF 18-SEP-1995; 95WO-US11859.
XX	
XX	PR 22-SEP-1994; 94US-0310468.
PA	(MERI ) MERCK & CO INC.
XX	
PI	Hofmann KJ, Jansen KU;
DR	WPI; 1996-188437/19.
XX	
PT	DNA encoding human papillomavirus 6a - used to treat and in
PT	vaccines against human papillomavirus 6a disease, e.g. condyloma
PT	acuminata
PS	Claim 1; Fig 1A-G; 46pp; English.
XX	
XX	A DNA molecule (AAT13413) codes for human papillomavirus 6a (HPV6a),
CC	the predominant HPV subtype found in biopsies of condyloma acuminata
CC	(benign lesions of the respiratory and genital mucosa). It was
CC	isolated from DNA extracted from a vulva condyloma acuminatum lesion
CC	following screening with a probe (see AAT13414) complementary to the 3

CC end of the HPV subtype 6b L1 gene. The DNA can be used to develop  
CC vaccines against HPV6a, to treat conditions caused by HPV6a, for  
CC HPV serotyping, and to produce recombinant HPV6a polypeptides.

SQ Sequence 8010 BP; 2470 A; 1529 C; 1715 G; 2296 T; 0 other;

Query Match	100.08;	Score 20;	DB 17;	Length 8010;
Best Local Similarity	100.08;	Pred. No. 0.58;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 TGCCTAATTCGGTGTCTACCTG 20  
 |||||  
 Db 401 TGCCTAATTCGGTGTCTACCTG 420

RESULT 10  
 ID ABX12328/c  
 XX ABX12328 standard; DNA; 23 BP.  
 AC  
 XX ABX12328;  
 DT 10-MAY-2003 (first entry)  
 XX  
 DE Fluorescent multiplex HPV PCR assay primer #4.  
 XX  
 KW Human papillomavirus; HPV; fluorescent multiplex HPV PCR assay; primer;  
 XX PCR; ss; polymerase; fluorophore; fluorescence.  
 OS Human papillomavirus.  
 XX  
 WO2003019143-A2.  
 XX  
 PN 06-MAR-2003.  
 PD  
 XX 19-AUG-2002; 2002WO-US26964.  
 PF  
 XX 23-AUG-2001; 2001US-314383P.  
 PR  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX Jansen KU, Taddeo FU, Li W, Dicello AC;  
 PI WPI; 2003-312914/30.  
 DR  
 XX  
 PT Detecting the presence of a human papillomavirus subtype, using  
 PT multiple fluorophores, in a nucleic acid-containing sample, useful in  
 PT PCR-based assays for identifying HPV subtypes  
 XX  
 XX Claim 15; Page 39; 59pp; English.  
 XX  
 XX The invention relates to a method for detecting the presence of a human  
 CC papillomavirus (HPV) subtype in a nucleic acid-containing sample. The  
 CC method comprises amplifying the nucleic acid in the presence of a  
 CC polymerase and a plurality of oligonucleotide sets, allowing digestion of  
 CC each fluorescent probe during amplification to dissociate the fluorophore  
 CC from the quencher molecule, detecting a change of fluorescence upon  
 CC dissociation, and determining if the sample is positive for the HPV  
 CC subtype if a change of fluorescence is detected in at least two emission  
 CC maxima. Each oligonucleotide set mentioned above consists of a forward  
 CC discriminatory PCR primer hybridising to a first location of an HPV  
 CC subtype, a reverse discriminatory PCR primer hybridising to a second  
 CC location of the HPV subtype downstream of the first location, and a  
 CC fluorescent probe labelled with a quencher molecule and a fluorophore  
 CC which emits energy at a unique emission maxima, the probe hybridising to  
 CC a location of the HPV subtype between the first and the second locations,  
 CC where each oligonucleotide set specifically hybridises to a different HPV  
 CC amplicon derived from the same HPV subtype. The change of fluorescence  
 CC corresponds to the occurrence of nucleic acid amplification. The method  
 CC of the present invention is useful in PCR-based assays for detecting HPV  
 CC subtypes in clinical samples. This sequence represents a PCR primer used  
 CC in the fluorescent multiplex HPV PCR assay of the invention.  
 XX  
 XX Sequence 23 BP; 8 A; 5 C; 6 G; 4 T; 0 other;

Query Match 90.0%; Score 18; DB 25; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTAATCGGTGCTACCTG 20  
|||||  
DB 23 CTAATCGGTGCTACCTG 6

RESULT 11

AA529985  
ID AA529985 standard; DNA; 281 BP.  
XX  
AC AA529985;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human lung antigen genomic DNA #55.  
XX  
XX Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; neotropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasia;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-fertility; food additive.  
XX  
XX Homo sapiens.  
XX  
XX WO200155303-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01301.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0186874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0233968.  
PR 14-SEP-2000; 2000US-0233997.  
PR 14-SEP-2000; 2000US-0233998.  
PR 14-SEP-2000; 2000US-0233999.  
PR 14-SEP-2000; 2000US-0234000.  
PR 14-SEP-2000; 2000US-0234001.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.



PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0233397.  
 PR 14-SEP-2000; 2000US-0233398.  
 PR 14-SEP-2000; 2000US-0233399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0238935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0251989.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251988.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 PI WPI, 2001-476224/51.  
 DR  
 XX  
 PT Isolated polypeptide for treating, preventing and/or prognosing  
 PT disorders related to the respiratory system including respiratory  
 PT cancers and also for testing and detection e.g. diagnosis -  
 XX  
 PS Disclosure; SED ID No 928; 546pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human  
 CC respiratory antigens (AAU17685-AAU1973), and cDNA and genomic  
 CC sequences encoding for these polypeptides. The sequences of the  
 CC invention are useful for preventing, treating and/or prognosing  
 CC disorders related to the respiratory system including throat  
 CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),  
 CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of  
 CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences  
 CC of the invention are useful in gene therapy and antisense therapy.  
 CC AA528161-AA528764 represent genomic sequences encoding for novel  
 CC human respiratory antigens.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPRO  
 CC at ftp.wipro.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 281 BP; 83 A; 63 C; 26 G; 109 T; 0 other;  
 Query Match 79.0%; Score 15.8; DB 22; Length 281;  
 Best Local Similarity 89.5%; Pred. No. 60;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 TGCCTAATTCGGTGCCT 19  
 Db 102 TGCCTAATTCGGTGCCT 120  
 RESULT 13  
 AAS29864  
 ID AAS29864 standard; cDNA; 287 BP.  
 XX  
 AC AAS29864;  
 XX  
 DT 21-NOV-2001 (first entry)  
 XX  
 DE Human lung antigen cDNA polynucleotide #26.  
 XX  
 KW Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;  
 KW Chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
 KW antirheumatic; antiproliferative; cyostatic; cardiant; neuroprotective;  
 KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
 KW ophthalmological; vulnetary; gene therapy; autoimmune disease; neoplasm;  
 KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;  
 KW cerebrovascular disorder; nervous system disorder; bacterial infection;

KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-fertility; food additive.  
XX  
XX Homo sapiens.  
XX  
PN WO200155303-A2.  
XX  
XX 02-AUG-2001.  
PF  
XX 17-JAN-2001; 2001MO-US01301.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250191.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.

PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-457723/49.  
XX P-PSDB; AAU16577.  
XX  
XX Isolated polypeptide for treating, preventing and/or prognosing  
XX respiratory disorders related to the lung including lung cancers and  
XX also for testing and detection e.g. diagnosis -  
XX  
XX Claim 1; SEQ ID No 36; 507pp; English.  
XX  
XX Sequences AAS29839-AAS29930 represent cDNA molecules, which encode the  
XX lung antigen polypeptides of the invention. Lung antigen polypeptides and  
XX their associated polynucleotides are useful in the diagnosis, treatment  
XX and prevention of various types of disorders in e.g. humans, mice,  
XX rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological  
XX condition can be determined by detecting the presence or absence of a  
XX mutation in a lung antigen polynucleotide. The treatable disorders  
XX include autoimmune diseases such as rheumatoid arthritis,  
XX hyperproliferative disorders such as neoplasms of the breast or liver,  
XX cardiovascular disorders such as cardiac arrest, cerebrovascular  
XX disorders such as cerebral ischaemia, nervous system disorders such as  
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi,  
XX ocular disorders such as corneal infection, endocrine disorders such as  
XX premature labour and infertility, gastrointestinal disorders such as  
XX Crohn's disease, renal disorders such as glomerulonephritis and  
XX respiratory disorders such as asthma and pleurisy. The polypeptides can  
XX also be used to aid wound healing, to prevent skin aging due to sunburn,  
XX to maintain organs before transplantation, to regenerate tissues and in  
XX chemotaxis. The polypeptides can also be used as a food additive or  
XX preservative to increase or decrease storage capabilities.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO

Query Match 79.0%; Score 15.8; DB 22; Length 287;  
Best Local Similarity 89.5%; Pred No. 80;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCATATTCGTGCTACCT 19  
||||| ||| ||||| |||  
DB 102 TGCATATTCGTGCTACCT 120

RESULT 14  
AAS28028  
ID AAS28028 standard; CDNA; 287 BP.  
XX  
XX AAS28028;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX Novel cDNA encoding for human respiratory antigen #160.  
XX  
XX Human; respiratory antigen; respiratory disorder; throat disorder;  
XX lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;  
XX anti allergic; anti asthmatic; anti inflammatory; olfactory;  
XX respiratory active; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200155448-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01333.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR

PR 04-FEB-2000; 2000US-0180628.  
PR 04-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198122.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.



CC progress of fermentation and the growth conditions of a fungus,  
 CC especially of *Aspergillus oryzae* which is widely used in industrial  
 CC fermentation. Also monitoring for fungal contamination.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

SQ Sequence 907 BP; 249 A; 208 C; 205 G; 243 T; 2 other;

Query Match 79.0%; Score 15.8; DB 25; Length 907;

Best Local Similarity 89.5%; Pred No. 91;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTCTACT 19

DB 154 TGCTTCTTCGGTGTCTACT 136

Search completed: August 23, 2003, 10:38:19  
 Job time : 208.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 08:28:23 ; Search time 993 Seconds  
(without alignments)  
823.960 Million cell updates/sec

Title: US-09-927-585A-1

Perfect score: 20  
Sequence: 1 tgcataaacctgtgtcc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_ov:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rnd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	20	100.0	20	6	AX301234	AX301234 Sequence
2	20	100.0	20	6	E05257	E05257 Primer for
3	18.4	92.0	20	6	E05247	E05247 Part of DNA
4	18.4	92.0	803	14	PPHE67	D10597 Human papill
5	18.4	92.0	7851	14	PPH35CG	M74117 Human papill
6	18.4	92.0	7879	14	HPV35H	X74477 Human papill
7	18.4	92.0	7912	14	PPH31A	J04353 Human papill
8	17.4	87.0	10999	1	AE013763	AE013763 Yersinia
9	17.4	87.0	102178	8	AP003803	AP003803 Oryza sat
10	17.4	87.0	223885	2	AC107604	AC107604 Rattus no
11	17.4	87.0	236021	2	AC094509	AC094509 Rattus no
12	17.4	87.0	236062	2	AC105688	AC105688 Rattus no
13	17.4	87.0	258050	1	AJ414153	AJ414153 Yersinia
14	16.8	84.0	20	6	E05251	E05251 Part of DNA
15	16.8	84.0	450	14	AF234529	AF234529 Human pap
16	16.8	84.0	450	14	AF234530	AF234530 Human pap
17	16.8	84.0	450	14	AF234531	AF234531 Human pap
18	16.8	84.0	450	14	AF478150	AF478150 Human pap
19	16.8	84.0	450	14	AF478151	AF478151 Human pap
20	16.8	84.0	450	14	AF478152	AF478152 Human pap
21	16.8	84.0	450	14	AF478153	AF478153 Human pap
22	16.8	84.0	450	14	AF478154	AF478154 Human pap
23	16.8	84.0	450	14	AF478155	AF478155 Human pap
24	16.8	84.0	450	14	AF478156	AF478156 Human pap
25	16.8	84.0	450	14	AF478157	AF478157 Human pap
26	16.8	84.0	450	14	AF478158	AF478158 Human pap
27	16.8	84.0	450	14	AF478159	AF478159 Human pap
28	16.8	84.0	450	14	AF478160	AF478160 Human pap
29	16.8	84.0	450	14	AF478161	AF478161 Human pap
30	16.8	84.0	450	14	AF478162	AF478162 Human pap
31	16.8	84.0	450	14	AF478163	AF478163 Human pap
32	16.8	84.0	450	14	AF478164	AF478164 Human pap
33	16.8	84.0	450	14	AF478165	AF478165 Human pap
34	16.8	84.0	450	14	AF478166	AF478166 Human pap
35	16.8	84.0	450	14	AF478167	AF478167 Human pap
36	16.8	84.0	766	6	AX186489	AX186489 Sequence
37	16.8	84.0	3443	6	AX536430	AX536430 Sequence
38	16.8	84.0	3840	8	YSCABOI	L31406 Saccharomyc
39	16.8	84.0	4569	8	SCYBL085W	Z3846 S.cerevisia
40	16.8	84.0	7801	14	D21208	D21208 Human papill
41	16.8	84.0	7812	14	AF131950	AF131950 Human pap
42	16.8	84.0	7824	14	PPH58	D90400 Human papill
43	16.8	84.0	7909	6	A07020	A07020 Human papill
44	16.8	84.0	7909	6	A12360	A12360 complete nu
45	16.8	84.0	7909	14	PPH33CG	M12732 Human papill

## ALIGNMENTS

RESULT 1  
AX301234  
LOCUS AX301234 20 bp DNA linear PAT 30-NOV-2001  
DEFINITION Sequence 3 from Patent WO0185994.  
ACCESSION AX301234  
VERSION AX301234.1 GI:17382322  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
1 Perseu, S., de Montis, A. and Floris, M.M.  
Method and means for identifying hpv virus  
Patent: WO 0185994-A 3 15-NOV-2001;  
Bioanalisi Centro Sud S.N.C. Di Perseu Sinbiado EC. (IT)

FEATURES  
Source

Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="primer representative of region E6 of strains of  
HPV virus having middle-high oncogenic hazard"

BASE COUNT  
5 a 5 c 4 g 6 t

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTTGTGTC 20  
1 TGTCAAAACCGTTGTGTC 20

## RESULT 2

E05257  
LOCUS E05257 20 bp DNA linear PAT 29-SEP-1997  
DEFINITION Primer for amplifying Human papillomavirus gene.  
ACCESSION E05257  
VERSION E05257.1 GI:2173447  
KEYWORDS JP 1993192200-A/27.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.  
TITLE DETECTION OF HUMAN PAPILLOMA VIRUS  
JOURNAL Patent: JP 1993192200-A 27 03-AUG-1993;  
TAKARA SHUZO CO LTD  
OS Artificial gene  
NC Artificial sequence; Genes.  
PN JP 1993192200-A/27  
PD 03-AUG-1993  
PF 19-AUG-1991 JP 1991230839  
PR 20-AUG-1990 JP 90P 217067  
PI OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI  
FUKUSHIMA MICHIO,  
FUJINAGA KEI  
PC C1201/70,C1201/68;  
CC strandedness: Single;  
CC topology: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

FEATURES  
Source

BASE COUNT  
5 a 5 c 4 g 6 t

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTTGTGTC 20  
1 TGTCAAAACCGTTGTGTC 20

## RESULT 3

E05247  
LOCUS E05247 20 bp DNA linear PAT 29-SEP-1997  
DEFINITION Part of DNA sequence of Human papillomavirus 31.  
ACCESSION E05247  
VERSION E05247.1 GI:2173437  
KEYWORDS JP 1993192200-A/17.  
SOURCE synthetic construct

ORGANISM  
synthetic construct

artificial sequences.  
1 (bases 1 to 20)  
Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.  
TITLE DETECTION OF HUMAN PAPILLOMA VIRUS  
JOURNAL Patent: JP 1993192200-A 17 03-AUG-1993;  
TAKARA SHUZO CO LTD  
OS Artificial gene  
NC Artificial sequence; Genes.  
PN JP 1993192200-A/17  
PD 03-AUG-1993  
PF 19-AUG-1991 JP 1991230839  
PR 20-AUG-1990 JP 90P 217067  
PI OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI  
FUKUSHIMA MICHIO,  
FUJINAGA KEI  
PC C1201/70,C1201/68;  
CC strandedness: Single;  
CC topology: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

FEATURES  
Source

BASE COUNT  
4 a 5 c 5 g 6 t

Query Match 92.0%; Score 18.4; DB 6; Length 20;  
Best Local Similarity 95.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTTGTGTC 20  
1 TGTCAAAACCGTTGTGTC 20

## RESULT 4

E05247  
LOCUS E05247 803 bp DNA linear VRL 02-FEB-1999  
DEFINITION Human papillomavirus gene for E6 and E7 proteins, complete cds.  
ACCESSION D10597  
VERSION D10597.1 GI:535788  
KEYWORDS E6 protein; E7 protein.  
SOURCE Human papillomavirus  
ORGANISM Human papillomavirus  
Virus(es); dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
1 (bases 1 to 803)  
Fujinaga,K.  
Unpublished  
REFERENCE 2 (bases 1 to 803)  
Fujinaga,K.  
Direct Submission  
Submitted (26-FEB-1992) Kei Fujinaga, Cancer Research Institute,  
Sapporo Medical College, Department of Molecular Biology, S1, W17,  
Chuo-ku, Sapporo 060, Japan (Tel:011-611-2111, Fax:011-615-3099)  
Location/Qualifiers  
1..803  
/organism="Human papillomavirus"  
/mol\_type="genomic DNA"  
/isolate="Type Sapporo-A"  
/db\_xref="taxon:10566"

FEATURES  
Source

## CDS

52..501  
/codon\_start=1  
/product="E6 protein"  
/protein\_id="BA01447.1"  
/db\_xref="GI:535788"  
/translation="MNPDAERPYLHDLCEVESIHEICINCVCYKQELQREYVD  
FACVDLCIYRRCQPGVCMKICFKYKISSEYRKYRYSYGETLEKCCNQGLLLLR  
CITCCPFLCPVERKHLEKRFNHNIGRGWTCGMCWCKPFRREYV"

```
CDS
504..803
/codon_start=1
/product="E7 protein"
/protein_id="BA01448.1"
/db_xref="GI:335790"
/translation="MHGEITTLQDYVLDEPEATDLYCYEOLDSSEEDRTIDPAG
OAKPDSNNIYTSCKCEATRLCQVSHIDIRKLEDLMTGTFGIVCGCSQRA"
BASE COUNT      282 a      127 c      190 g      204 t
ORIGIN

Query Match      92.0%; Score 18.4; DB 14; Length 803;
Best Local Similarity 95.0%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TGTCAAAACCGTGTGTC 20
Db      367 TGTCAAAACCGTGTGTC 386

RESULT 5
PPH35CG      7851 bp      DNA      circular VRL 10-MAY-2002
LOCUS      Human papillomavirus type 35 complete genome.
DEFINITION
ACCESSION      M74117
VERSION      M74117.1 GI:333050
KEYWORDS      complete genome; major capsid protein; minor capsid protein;
SOURCE      Human papillomavirus type 35
ORGANISM      Human papillomavirus type 35
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
1 (sites)
REFERENCE      1 Lorincz,A.T., Quinn,A.P., Lancaster,W.D. and Temple,G.F.
AUTHORS      A New Type of Papillomavirus Associated with Cancer of the Uterine
TITLE      Cervix
JOURNAL      Virology 159, 187-190 (1991)
MEDLINE      87265470
REFERENCE      2 (bases 1 to 7851)
AUTHORS      Marich,J.E., Ponteler,A.V., Rice,S.M., McGraw,K.A. and
TITLE      Dubensky,T.W.
JOURNAL      The Phylogenetic relationship and complete nucleotide sequence of
MEDLINE      Virology 186 (2), 770-776 (1992)
PUBMED      92124753
COMMENT      Original source text: Human papillomavirus type 35 cervical
carcinoma DNA.
FEATURES
source      Location/Qualifiers
1..7851
/organism="Human papillomavirus type 35"
/proviral
/mol_type="genomic DNA"
/isolation_source="cervical carcinoma"
/db_xref="taxon:10587"
join(7092..7851,1..109)
/standard_name="tCR"
/function="regulatory region"
17..22
/bound_moiety="Sp-1"
complement(18..26)
24..35
/note="putative"
/bound_moiety="E2"
/function="gene transcription"
39..50
/note="putative"
/bound_moiety="E2"
/function="gene transcription"
54..59
/note="putative"
/function="gene transcription"
54..59
/note="putative"
TATA_signal
gene
110..559
/gene="E6"
110..559
CDS

/misc_feature
/misc_feature
gene
CDS
gene
CDS
/gene="E6"
/note="putative: E6 ORF from bp 59 to 559"
/codon_start=1
/product="transforming protein"
/protein_id="AAA4696.1"
/db_xref="GI:333051"
/translation="MFQDAEPERYKRLHDCNEVESIHEICLNCYCKOELORSEYD
FACDYCIYRREGOPYGVCMCKLRFKPSISEYRWRYSVYGETLEKCKKOLCHLLIR
CITCKPKPCPEKORHLEKKRRHNIIGRNTGCMCKMPTRETEY"
230..238
/gene="E6"
/standard_name="Splice donor"
405..416
/gene="E6"
/standard_name="Splice acceptor"
562..861
/gene="E7"
562..861
/gene="E7"
/note="putative"
/codon_start=1
/product="transforming protein"
/protein_id="AAA46967.1"
/db_xref="GI:333052"
/translation="MHGEITTLQDYVLDEPEATDLYCYEOLDSSEEDRTIDPAG
OAKPDSNNIYTSCKCEATRLCQVSHIDIRKLEDLMTGTFGIVCGCSQRA"
868..2760
/gene="E1"
868..2760
/gene="E1"
/note="putative"
/codon_start=1
/product="replication protein"
/protein_id="AAA46968.1"
/db_xref="GI:333053"
/translation="MADPAGTDEGEGTCGNGFVEAVVSRRTGSSVEDENEDCDRG
EDMVDIFNDTILNIOAETFOALPEHAGQETHEKAVGVLRKRYASPLSSVLCVN
NNISPEKRLACIENTKTAKRRLPEPSGYSNSEEIHEIQOVGHGDEGCSGSC
DSITSSDERRHETPRDILQLKSNMAAMALAKRELFGISFTELLRPPRSXSTC
TMCVAFAGIAPSVANFKHITTVYITNVRHGAANYTLALRKVKERKDOUKTIDAK
LLCISASMLIOPKLRSTPALYWKFTAMSNISEVDGETPMPIOQIVLOASHFNDAI
FLDSEVQWAVYNDPFDSDIDIAKYKQLAETNSNCAFLKSNQAKIVLQSCMCHY
KRAEKREMTSQMIRKCAQVDDGMDRDIYFELRYQOYDPAVAFSALKNPLHGVPK
NCLIIYGAPNTGKSLSGMSLMHPIQGLAII SYVNSKSHFWLOPLDYDAKIMLADATSPC
GIYRPIFKCTRKAKTSYISRCALSLIVHIMPTPTTYIINNAGCKDDWPLHSHVYFT
FHNPEPFDNNGNPEITGLNDKNMKSFPSTKWCRLNLHDEEVKENDGAPFAFCVSGN
TETLRD"
2693..3796
/gene="E2"
2693..3796
/gene="E2"
/function="regulation of gene expression"
/note="putative"
/codon_start=1
/product="regulatory protein"
/protein_id="AAA46969.1"
/db_xref="GI:333054"
/translation="MMETLSQRLSVQDKILBHYETDSTCLSDHIQYKILRECAVF
YKAREGIRTLNHOVPTQAIKAKAMQAIELQIMLETINTETCYSTEDWTQDETSEL
YTVVTPRCLKRDVYVEAOFDQDKQMTMHTYTWHTIYILSDSCITVAKGLVANKIY
VHQGVETYYVTFREBEAKKYGKKNIMVHGGGVYICPEVSFSTELSTAEIATQHLAY
NTTEHTKACSVGTETQKTNKRRLRGTELEYPNPKRYRLSAYDSVDGVYSTCCT
NKDRGSCSTTPIYHLKQDANTLCKSRRLCKYKALYODASTRWCTNCTNDRKQAI
VTLITTYTERQKRLTIVKIPNTIVVSKGIMSI"
3273..3563
/gene="E4"
3273..3563
/gene="E4"
/note="putative"
/codon_start=1
/product="unknown"
/protein_id="AAA46965.2"
/db_xref="GI:20522255"
```

gene  
/translation="MEVNLVLAQNVDLKLHLSYPTTPPPRRIIPAPAPAPQPKPR  
OITNFEAVSSPPTTPPECCDSVPMTVLTBESTLHLTAQRTGVVVVVQ.LH"  
3793..4038.  
/gene="E5"  
3793..4038  
/gene="E5"  
/note="putative"  
/codon\_start=1  
/product="E5"  
/protein\_id="AAA46970.1"  
/db\_xref="gi:333056"  
/translation="MIDLTASTVYLACFLCFCVLLCICLLVRSLLSVLSYSL.IHL  
VLILVAVATPPLAAVSSCFIYLMIMINAAQYLAVQ"  
4159..4164  
/note="putative"  
4184..5593  
/gene="L2"  
4184..5593  
/gene="L2"  
/note="putative; putative"  
/codon\_start=1  
/product="minor capsid protein"  
/protein\_id="AAA46971.1"  
/db\_xref="gi:333057"  
/translation="WRKRSTIKRVKASATQLYRTCKAAGTCPPDVLPKEGNTVADQ  
ILKYSMAVFEGLDIGSGSGTGRSGYPLGTPTPATNIPRPVYESI.DLDTI  
GTLDSIVLVEETSFIEISGAPVVPVPTGTGFIITSTPTALIDVTSISHDNBP  
PTTDSVLPPTPATPSGHFVLLSSSIETHVKEIEMPTFIVSDNNTSTPTIPSSP  
RPTYLGIKSKTGQOVKVDPAFMTSPAKLITLYDPAEGLNPTTLOEHEDISLAR  
DDEDDIILAHPRALTSRKRTIRYRANKRMTMRSCAKGAVHYODLSSTEDDI  
EQAPLOHPSSLPHTVYSTSLDNKRFDIYAPIDREDDIEFSASSNLTITTSNATVP  
STPTPLPSGDIPITGAPDIVFNSNTITNSVLPPTGPTISIIDGDEFYLAHSYVL  
LRRRKRAPIFVADSVAV"  
5574..7091  
/gene="L1"  
5574..7091  
/gene="L1"  
/gene="L1"  
/note="putative"  
/codon\_start=1  
/product="major capsid protein"  
/protein\_id="AAA46972.1"  
/db\_xref="gi:333058"  
/translation="MSLMRSNEATVYLLPVSYSKVYSDDEVYTRNTIYHAGSRLLAA  
VGHPIYAIKQDSNKLAVPKVSGLOYRFRVRLKIDPKNKEGPDPSFVDPCIALVMAQ  
LIEVVGROGPIGVGISGPHLLNKDDTLENLKKYVNSGNSGTDRRECISMYKQTQLC  
LIGCPRLDIEHNGKTPCPCANOVKAGECPLELMTV.LQDGMVDITGGANDFTLQ  
NNSVDPLDCCSICKYPPYLKAVSEPYDMLFEYLRBQMFVRHLFNKAGVGGTTPA  
DIYIKGTGLTLPSTSYFPTPGSAMVTSQAQIFNKYWLQRAQHNNIGCWNQTLV  
VPTSTNNKSVCSAVSSDSYTKNDNFKEYELRHGEEDYDQIFQLCITLADYATVYI  
HMSNISLIDNMFGLTPPGSGTLEDYTRYVTSVAVTCOKPSAPKRDPLPLNFEMWE  
DKERKPSLIDNDFPLGRKRLFLQAGLKAHPNRILGRAPASTSKKSSYTKRRKVAS"  
7090..7105  
7123..7160  
7415..7426  
/note="putative"  
/bound\_moiety="E2"  
/function="gene transcription"  
7477..7491  
/standard\_name="glucocorticoid responsive element"  
/bound\_moiety="hormone receptor"  
7514..7519  
/bound\_moiety="NF-1"  
7527..7532  
/bound\_moiety="NF-1"  
7534..7539  
/bound\_moiety="NF-1"  
/complement(7670..7677)  
/standard\_name="keratinocyte specific enhancer  
(CK-ocloner)"  
/note="putative"  
7671..7676  
/bound\_moiety="NF-1"  
7695..7700

BASE COUNT	2553 a	/bound_molecy="NF-1"	1343 c	1568 g	2387 t
Query Match	92.0%	Score 18.4;	DB 14;	Length 7851;	
Best Local Similarity	95.0%;	Pred. No. 43;			
Matches	19; Conservative	0; Mismatches	1;	Indels	0; Gaps
OY	1	TGTCAAAACCGTTGTCGC	20		
Db	425	TGTCAAAACCGCTGTGTCC	444		
RESULT 6					
HPV35H					
LOCUS	Human papillomavirus type 35H	genomic DNA.	Linear	VRL 24-FEB-1999	
DEFINITION	X74477.1	GI:396997			
ACCESSION	E1 gene; E2 gene; E4 gene; E6 gene; E7 gene; early protein; L1				
VERSION	gene; L2 gene; late protein.				
KEYWORDS	Human papillomavirus type 35H				
SOURCE	Human papillomavirus type 35H				
ORGANISM	viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				
REFERENCE	Papillomavirus.				
AUTHORS	1 (bases 1 to 7879)				
TITLE	Dellius, H. and Hofmann, B.				
JOURNAL	Primer-directed sequencing of human papillomavirus types				
MEDLINE	Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)				
PUBMED	94265501				
REFERENCE	8205838				
AUTHORS	2 (bases 1 to 7879)				
TITLE	Dellius, H.				
JOURNAL	Direct Submission				
MEDLINE	Submitted (06-AUG-1993) H. Dellius, Deutsches				
PUBMED	Krebsforschungszenntum, Abteilung ATV, Im Neuenheimer Feld 506, W				
FEATURES	6900 Heidelberg, FRG				
Source	Location/Qualifiers				
	1..7879				
	/organism="human papillomavirus type 35H"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:31551"				
	/clone="insert in BamHI site of pBR322"				
	110..559				
	/gene="E6"				
	110..559				
	/gene="E6"				
	/codon_start=1				
	/product="early protein"				
	/protein_id="CA52561.1"				
	/db_xref="GI:396998"				
	/db_xref="SWISS-PROT:P27228"				
	/translation="MFDPARPKYKIDHCVEFSTHEICINVCVKQELORSEVD				
	FATCDLCLVREGQPIGVCMKCLFYKSISRYRTRYSVGTELEKCKNQLCHLLIR				
	CITCQKPCLPVEKQRHLLEKKRFHNIGRWGMSCWKPRRETV"				
	562..861				
	/gene="E7"				
	562..861				
	/gene="E7"				
	/codon_start=1				
	/product="early protein"				
	/protein_id="CA52562.1"				
	/db_xref="GI:396999"				
	/db_xref="SWISS-PROT:P27230"				
	/translation="MHGITLIDLYVIDLEAFANDIYCYEQLCDSSSEEPDITDPAG				
	QAKPTSNNTIVTSCKGEATRLRCVOSTHIDIKLEDLMGTGVIYCPCGSQA"				
	868..2781				
	/gene="E1"				
	868..2781				
	/gene="E1"				
	/codon_start=1				
	/product="early protein"				
	/protein_id="CA52563.1"				
	/db_xref="GI:397000"				

Query Match	92.0%	Score 18.4	DB 14	Length 7879
Best Local Similarity	95.0%	Pred. No. 43		
Matches 19	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	TGTCAAAACCGGTGTGTC	20	
Db	425	TGTCAAAACCGGTGTGTC	444	
RESULT 7				
PPH31A				

```

/protein_id="AA46982.1"
/db_xref="GI:459918"
/translation="MADPAGTDGRTGTCNGMFVEYVINDIQTGDISDENBDSQYD
EDMADFDLNCENVYNNQAEATQALPFAQZCEAEAEVQVLRKRYGSPDLSISGCV
INSPRKALICIENNSTAKRLRLEPDLSTIGTVEVQDQVVEBOQLTSCSGSG
THSERENETPTNRNLLOVLRKTSNGKAALGFKELGYSPNELLRPQSNRSTGTDWCY
AAGVGTETVAEGERTLLQPCYLCYCHSLGSGWGMALMVRRCANRSTITIKLERL
LLCISTMCLIQPRKLRTSTAALYKATVLTGNSIDVYGEIPEPVEERTVYLAQSHENDT
FDLSOMVQWAVYDNDVDDSEIAYKYATLQSDNSACAFKLSNSQAKLVKQCSGCRFY
KRKREKQMSMGOMIKSRCDRVSDSGEDDRIIVKFLRFOQIVFVSLATLKLGVKPR
NCLILHAPVTMGKRSFGMSLSLIFQGCQIIVYANSKSRFQVPLADATIGMLDPTFC

```

gene  
CDS  
2693..3811  
/gene="E2"  
/note="ORF E2 from bp 2663 to 3811"  
/codon\_start=1  
/product="regulatory protein"  
/protein\_id="AAA46953.1"  
/db\_xref="GI:459919"  
/translation="METLSORLNCQDKILEHYNDKRCDCDHIDYKHLRLCEVLMY  
KARMGJHSINHOVPALSVSKAKALQELQMLLETNLTRYKNEEDTMQOTSLEY  
LTAPGCKLKHGYVEVDFGDVHNTMTNMFYICIDGCTVEGQVCKGIYV  
HEGITYEVNTEAKKYGCTGKMEVHAGGVIVPESEVSDISFAGIYTKLPTAN  
NTTSNKTCALGTESEGRBATSTKRPREPERHNRHPKLLRGSDVDSVNGGIS  
AAACTNOTRANSCPATPTPIHLKGDANILKRLRSLKQLYEQVSSTWMTCTDCK  
HKNAIVTLTYISTSORDFLNTVKIPNTVSVSTGYMTI"  
3270..3578  
/gene="E4"  
/note="ORF E4 from bp 3270 to 3578"  
/codon\_start=1  
/protein\_id="AAA46949.1"  
/db\_xref="GI:459915"  
/translation="LEFLNLYAVTKYKYLGLDSYOQPTTPPHRIPKAPAPVAVK  
GRRRLSDQSGOSTEPPTPTSCCATPTVSTVGLQHLNQTQGLSVVQLHL"  
gene  
CDS  
3816..4070  
/gene="E5"  
/note="ORF E5 from bp 3804 to 4070"  
/codon\_start=1  
/protein\_id="AAA46954.1"  
/db\_xref="GI:459920"  
/translation="MIELNISTVSVILFCFLFCVCLVIRPLVLSVVALLL  
LIVILWVATSPRLRCFCYVYFVIFLVFHTHASFQSQ"  
4099..4134  
/note="putative"  
/gene="E2"  
/note="ORF L2 from bp 4060 to bp 5571"  
/codon\_start=1  
/product="minor capsid protein"  
/protein\_id="AAA46955.1"  
/db\_xref="GI:459921"  
/translation="MRSKRSPKTRKRASATOLYOTCKRAGCPSPDVIPKIEHTIADQ  
ILRSGMVGFCGICGSGCTGTCVPTSTRPTSEASIPRPRVSDIDPRGPD  
PSYLVESGIVVGAAPAPIPHPPTSGFATATADTAPILVDTVSSTENPEFD  
PSVLQPPAPATSGHLLISSSISTHNEELPMQFIVASTNNETSTPLPGARRA  
RLGYSKATQGVKVIDPFLSAPQOLITENPATEVAEBSLFSNTSHNADPDR  
LDIALHPALTSRRNTYRSLGNKOTLRTSGATIGARHYVYDISINPAGSIE  
MOPGASATSTSLNDGLYDADTDFVDPRATHNVSPTAVQSTSAVYPTNT  
VPLSTGPDPIFSGPDVIEHAPVQVPPFLAPTTPOVSTFVDSGDFLHNSYMLAK  
RRKRVSTFTDVAA"  
5552..7066  
/gene="L1"  
/note="ORF L1 from bp 5516 to 7066"  
/codon\_start=1  
/product="major capsid protein"  
/protein\_id="AAA46956.1"  
/db\_xref="GI:459922"  
/translation="MSLMRPSSEATVYLDPVPVSVSVNDEVTETNTIYHAGSARILTL  
VGHPYISIPKSDNPKIYVPRVSGIQRVFRVRLDPNKKFGFPDTSFTNPETQRLVMA

CVCLEVGRCPLGVGIGCHPLNKFEDTENSNRAGCPGTDNRCISMDYKOTOLCLL  
GCKPPIGEHMGKSPSCSNMATTEDCPLELTKNSVIDGDMVDFGAMDFATADTK  
SNVPLDICSICKPYDKMAYVYGTLEFFYLRDEMFRVHFNRSIGTAGEVPTDL  
YIKGSGSTATLANSTYPTPPSGSKVATSDAQLFNRPYMOAQAQNNICMGNOLETVY  
VDTRSTNMSVCAIANSDDPTFKSNKEVIRHGEEDLOFIPOLCKITTSADIMETI  
HSMNPALIDEMNRELTTPPGSGLEDTRYPTISOAITQKTRAPCKRPDKDYEMVEV  
MLKEFSADLDQFPGLGRKFLQAGTRAPKFRKAKRSAPSTTPAKRKTKK"  
7227..7231  
/note="putative"  
/standard\_name="glucocorticoid responsive element"  
/bound\_moiety="hormone receptor"  
7477..7488  
/bound\_moiety="E2"  
/function="gene transcription"  
7542..7549  
/standard\_name="keratinocyte-dependant enhancer"  
7868..7879  
/bound\_moiety="E2"  
/function="gene transcription"  
BASE COUNT 2528 a 1364 c 1572 g 2448 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 14; Length 7912;  
Best Local Similarity 95.0%; Pred. No. 43;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TGTCAAAACCGTTGTGTC 20  
||||| |||||||  
Db 423 TGTCAAAGACCGTTGTGTC 442  
RESULT 8  
AE013763 10909 bp DNA linear BCT 26-JUL-2002  
LOCUS Yersinia pestis KIM section 163 of 415 of the complete genome.  
DEFINITION AE013763 AE009952  
ACCESSION AE013763.1 GI:21958388  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Yersinia pestis KIM  
Yersinia pestis KIM  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Yersinia.  
REFERENCE  
1 (bases 1 to 10909)  
Deng, W., Burland, V., Plunkett, G., III, Boutin, A., Mayhew, G. F.,  
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,  
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,  
Straley, S. C., McDonough, K. A., Nilles, M. L., Watson, J. S.,  
Blattner, F. R. and Perry, R. D.  
Genome Sequence of Yersinia pestis KIM  
J. Bacteriol. 184 (16), 4601-4611 (2002)  
12142430  
2 (bases 1 to 10909)  
Deng, W., Burland, V., Plunkett, G., III, Boutin, A., Mayhew, G. F.,  
Liss, P., Perna, N. T., Rose, D. J., Mau, B., Zhou, S., Schwartz, D. C.,  
Fetherston, J. D., Lindler, L. E., Brubaker, R. R., Plana, G. V.,  
Straley, S. C., McDonough, K. A., Nilles, M. L., Watson, J. S.,  
Blattner, F. R. and Perry, R. D.  
Direct Submission  
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445  
Henry Mall, Madison, WI 53706, USA  
LOCATION/Qualifiers  
1..10909  
/organism="Yersinia pestis KIM"  
/mol\_type="genomic DNA"  
/strain="KIM"  
/db\_xref="taxon:187410"  
116..1126  
/gene="usg"  
/note="synonym: y1598"

CDS  
116..1126  
/gene="usg"  
/function="putative"  
/note="contains semialdehyde dehydrogenase domains;  
residues 1 to 336 of 336 are 71.81 pct identical to  
residues 1 to 337 of 337 from E. coli K12 : B2317;  
residues 1 to 336 of 336 are 71.51 pct identical to  
residues 1 to 337 of 337 from GenPept : >gb|AL21267.1|  
(AE008806) putative aspartate-semialdehyde dehydrogenase  
[Salmonella typhimurium LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="putative enzyme"  
/protein\_id="AA085167.1"  
/db\_xref="GI:21958389"  
/translation="MSDGNRIALGTCGAVGALLLELNERQPPVELYPLASERAG  
ATVRNGKSLIVENVAEFDSQALFVAGRASQVAAEAGNACLVYDSSGLFAM  
EPDPLVPSVNPQVADYRNRIYVADSLTSQLTIAKPLTEQAGLRHLVNTLS  
ASAHGKAADVADLAGOSAKLNLGIPAEQGFEPKQALFNVLPLADEGYSSEERLVDO  
VRKVLDEGLPTVSCIOSPFYGHQVHLEALRPIAAEARSLENCDDIOLSEED  
DYPTQSDASGSALSIGCVRNDYGIPEVLQFMSVADNIRFGALMAIQTAERLLQEO  
LY"  
gene  
1126..2007  
/gene="trua"  
/note="synonym: y1599"  
1126..2007  
/gene="trua"  
/function="enzyme; aminoacyl tRNA synthetases, tRNA  
modification  
/note="residues 20 to 290 of 293 are 78.22 pct identical  
to residues 4 to 270 of 270 from E. coli K12 : B2318;  
residues 20 to 290 of 293 are 78.59 pct identical to  
residues 4 to 270 of 270 from GenPept : >emb|CAD07600.1|  
(AL62274) tRNA pseudouridine synthase A [Salmonella  
enterica subsp. enterica serovar typhii]"  
/codon\_start=1  
/transl\_table=11  
/product="pseudouridylylate synthase I"  
/protein\_id="AA085168.1"  
/db\_xref="GI:21958390"  
/translation="MSGFELDDEBTLVQONTLAQOEKPAEVAGLKTALDIEYNGSRF  
GWRQEVASVOACLEALSKVANEPIGVFCAGRTDAGVHATGOVHPTSAVRKDA  
WMTGVNSHLPADIAVRWVTVDNFHARSATARRRYIIFSHRYPAIIAOGVTICY  
MPLDAEKMERAAOCLLGENDFSEFRAVOCSPARMNVAKHVTRGAVIYVDIKANA  
FVHHWRNVIGSLIEICGCGNDVTMAELLALKDRRAATAKADSLYLVSDYDPHF  
ALPKVPMGLPLAADBS"  
2137..2826  
/gene="deda"  
/note="synonym: y1600"  
2137..2826  
/gene="deda"  
/note="residues 10 to 225 of 229 are 85.64 pct identical  
to residues 1 to 216 of 219 from E. coli K12 : B2317;  
residues 10 to 225 of 229 are 85.18 pct identical to  
residues 1 to 216 of 219 from GenPept : >gb|AL21268.1|  
(AE008806) putative Deda family [Salmonella typhimurium  
LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA085169.1"  
/db\_xref="GI:21958391"  
/translation="MCLFISRVFMERFIRVDFITINDIVHAEVLAQGIWYALIEFL  
ILFCETGLVVPFLPGDSLFLVAGALASPLNDLVHTVTLMAVLAALIGDAINTTIG  
RIFGKRLFSNPDSIFRSYLDKTHQFYEHGKGAIIILARFPIIRTEFAPFVAGGKM  
SYRHPAAVYIGALVWVLLFTYAGYLFQGNVPIVQNNKLKLIYIIVSYILPGVEVWR  
HRRAAVRQNO"  
2901..3971  
/gene="accd"  
/note="synonym: y1601"  
2901..3971  
/gene="accd"  
/function="enzyme; fatty acid and phosphatidic acid

gene  
biosynthesis"  
/note="residues 53 to 350 of 356 are 86.24 pct identical  
to residues 1 to 296 of 304 from E. coli K12 : B2316;  
residues 53 to 350 of 356 are 86.91 pct identical to  
residues 1 to 296 of 304 from GenPept : >gb|AL21267.1|  
(AE008806) acetylCoA carboxylase, beta subunit [Salmonella  
typhimurium LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="acetyl CoA carboxylase, carboxyltransferase  
component, beta subunit"  
/protein\_id="AA085170.1"  
/db\_xref="GI:21958392"  
/translation="MWSVPADNPSPGALATSHWRVSAFFIGFSGRWLKHQRGNS  
QVOTERSMSNTERILNKSNTITOTKRASTPEGVWYTCDSGGVLYLAERLNEVYK  
COHHRMSAKARLHMLLDGSEVEIASLEPDIILFRSKRTDRIISAQKDTGERK  
ALVAMKGTQGMPTVAASFEAFMGSMASVVGARVAVAEALENCPLVCFSSGG  
ARMOEALMSLMQAKTSALAKOERGLPIEVLDPDTMGQVSAASLAMLGINIAEPK  
ALIGFSPVIEQTVREKLPFGFQSEFLIERGALDIMIVRRVMROTLASILSKLTHQ  
PPSVYVESKADTVQAQENQADV"  
4142..5524  
/gene="folC"  
/note="synonym: y1602"  
4142..5524  
/gene="folC"  
/function="enzyme; biosynthesis of cofactors, carriers:  
folic acid"  
/note="dihydrofolate synthetase; residues 31 to 458 of 460  
are 65.42 pct identical to residues 5 to 414 of 422 from  
E. coli K12 : B2315; residues 27 to 458 of 460 are 65.20  
pct identical to residues 1 to 414 of 422 from GenPept :  
>gb|AL21266.1| (AE008806) multifunctional  
folylpolyglutamate synthase; dihydrofolate synthase, also  
has formyltHP polyglutamate synthase activity [Salmonella  
typhimurium LT2]"  
/codon\_start=1  
/transl\_table=11  
/product="dihydrofolate:folylpolyglutamate synthetase"  
/protein\_id="AA085171.1"  
/db\_xref="GI:21958393"  
/translation="MGTTGASVQATWSIAGIKPREVTGLMNNHOTPQATSPLAALC  
YLEHLSPDIEGLERKQVABERLDLKPAPKIFVAGNNGTGTCCTEALILAGL  
RNGVSSPILKATREVRIOGGLSEASHSFAQIEARGDGLSYFEGFTSLALF  
PROAKDDVILTEVGGRLDANINYSVAATSLDHTDMLGIVREGSIRGKACV  
RGCKRAVGEPPMPOSINDVAEFGQLRRVAVKFSQOEPDQOEPDQOINGNH  
QCGEROLGCLPVNPLANAATLALVHTSELPLSDEATROGLQASLSGRQVYSQ  
PLLILDVANPHARVLRVLAOVINPVASRQGVRAVAGMLSDIAGTACLSER  
VDEWYCAPLEGRCASAGQLAEHLVSARQFSVETAMROMADADTDQDVIVCGSFHT  
VAHVMAALHL"  
5638..6366  
/gene="dedd"  
/note="synonym: y1603"  
5638..6366  
/gene="dedd"  
/function="putative membrane"  
/note="residues 12 to 239 of 242 are 55.45 pct identical  
to residues 1 to 209 of 211 from E. coli K12 : B2314;  
residues 3 to 239 of 242 are 56.72 pct identical to  
residues 1 to 218 of 220 from GenPept :  
>gb|AAG57443.1|AE005463.6 (AE005463) putative lipoprotein  
[Escherichia coli O157:H7 EDL933]"  
/codon\_start=1  
/transl\_table=11  
/product="putative lipoprotein"  
/protein\_id="AA085172.1"  
/db\_xref="GI:21958394"  
/translation="MRVASQFQNRVLTGIIIVAGIVYVGLLDGKKKHDEDEFAIP  
LYPKRDSOEIDIVPPLNDPLITTPREGAAQVAGLENDVSOAANDAGNLAVEPTI  
VAPPTQSKVEVFKPTVTKPVQKPLETKPIETKPAEKRPLEKRPKEKPEKAP  
VGOAVVOIGALKNAKNAKVEIYAMRLSGHRAFTVPATPVQDEITRLVYVGPASRQKL  
OSALPELNASLGINQVKKPYTTS"  
6663..7172  
/gene="cvpa"

CDS

/note="synonym: y1604"  
6663..7172  
/gene="cvpa"  
/function="membrane; colicin-related functions"  
/note="required for colicin V production; residues 1 to 160 of 169 are 83.12 pct identical to residues 1 to 160 of 162 from E. coli K12; residues 1 to 160 of 169 are 83.12 pct identical to residues 1 to 160 of 162 from Genpept: >gb|AL21264.1| (AE008806) membrane protein required for colicin V production [Salmonella typhimurium LT2]"  
/codon\_start=1  
/transl\_table=1  
/product="membrane protein"  
/protein\_id="AAM85173.1"  
/db\_xref="GI:21958395"  
/translation="MWMIDYIIGIGSAISVLRGVRREALSLYMGCAFVASH  
YTYAVYTRFEDETVRNGIAIGLIFATLIVGAIIVYVSSIVERTGLSGTDRLVG

Query Match 87.0%; Score 17.4; DB 1; Length 10909;  
Best Local Similarity 94.7%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTCTC 19  
|||||  
8159 TGTCAAAACCGTATGTC 8177

RESULT 9  
AP003803/C 102178 bp DNA linear PLN 26-OCT-2002  
DEFINITION *Oryza sativa* (japonica cultivar-group) genomic DNA, chromosome 7,  
AP003803  
ACCESSION AP003803.3 GI:24414010  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group) Embryophyta; Tracheophyta;  
Eukaryota; Viridiplantae; Streptophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nippondare(GAS) genomic DNA, chromosome 7, BAC  
clone:OJ1060\_D03  
Published Only In Database (2001)  
2 (bases 1 to 102178)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (03-JUL-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@ias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Oct 26, 2002 this sequence version replaced gi:2275423.  
Genes were predicted from the integrated results of the following:  
GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH  
(http://www.softberry.com/), Genemark.hmm  
(http://opal.biology.gatech.edu/genemark/), Glimmer  
(http://www.tigr.org/tdb/glimmer/glmr form.html), RiceHMM  
(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor  
(http://bioinformatics.lastate.edu/cgi-bin/sp.cgi), BLASTN and  
BLASTX. The genomic sequence was searched against NCBI NonRedundant  
protein database, nr  
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
RGP. Protein homologues of the coding regions were searched against  
NCBI NonRedundant protein database with BLASTP. ESTs represent the  
identified cDNA sequences using BLASTN with the corresponding DBJ  
accession no. and RGP clone ID.  
A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with EST homology (covering

almost the entire length of partial sequence) is classified as an  
'unknown' protein. A gene predicted by two or more gene prediction  
programs is classified as a 'hypothetical' protein according to  
IRGSP standard. A gene predicted by a single gene prediction  
program is also classified as a probable 'hypothetical' protein and  
is included as a miscellaneous feature of the sequence.  
The orientation of the sequence is from -21M13 to M13rev of the BAC  
clone. This sequence of OJ1060\_D03 clone has an overlap with  
OJ1127\_E01 clone(DBJ: AP003747) at 5' end and an overlap with  
OJ1113\_E01 (DBJ: AP005437) at 3' end. The nucleotide sequence of  
this BAC clone was generated by combining Monsanto and RGP-Japan  
sequencing data. Detailed information on overlap and assembly  
quality together with annotation of this entry is available at  
http://rgp.dna.affrc.go.jp/Genomeseg.html.

FEATURES  
Source  
Location/Qualifiers  
1..102178  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nippondare"  
/db\_xref="taxon:39947"  
/chromosome="7"  
/clone="OJ1060\_D03"  
1311..1511  
/gene="OJ1060\_D03.1"  
1311..1511  
/gene="OJ1060\_D03.1"  
/note="This category is not included in IRGSP standard.  
hypothetical ORF  
predicted by GlimmerM"  
complement(4270..4578)  
/gene="OJ1060\_D03.2"  
complement(4270..4578)  
/gene="OJ1060\_D03.2"  
/note="hypothetical protein  
predicted by Genemark.hmm etc."  
/codon\_start=1  
/protein\_id="BAC2262.1"  
/db\_xref="GI:24414011"  
/translation="MNQTFKVRGLGLPTNCTAGHSLSVHLQLSMHTPTFPVY  
HPSILACDDELGPFLGASPPRLFLLAGADVVYVAGRAFSAGGRYERSAASWRS  
"  
complement(join(5870..6466,6570..6647))  
/gene="OJ1060\_D03.3"  
complement(join(5870..6466,6570..6647))  
/gene="OJ1060\_D03.3"  
/note="hypothetical protein  
predicted by Genemark.hmm etc."  
/codon\_start=1  
/protein\_id="BAC2263.1"  
/db\_xref="GI:24414012"  
/translation="MERRLTPGVRAWSLHRLNMNDAAELSVGRRLGCAVPCPEEY  
VVGDEMLHTISLTPPVLPPLSSLSLSLFFAGSGTGLVSGGAADLGEAP  
SSPSSLDGGLHAAPVAVSNDNDNDVSCMLKRLPLVCGSLIYINPVYGSLSH  
PSPPLATRSRSHAGMISFGYSTGTGKRIYVFPNGGLVDEVTLGRASSSSPPAMA  
ATWGW"  
join(8621..8754,8874..8983,9860..9930,10812..10932,  
11030..11337)  
/gene="OJ1060\_D03.4"  
join(8621..8754,8874..8983,9860..9930,10812..10932,  
11030..11337)  
/gene="OJ1060\_D03.4"  
/note="mutator-like transposase-like protein"  
/codon\_start=1  
/protein\_id="BAC2263.1"  
/db\_xref="GI:24414013"  
/translation="MRRRLHGMARCGASAGSSSPALOSGSPFASTLONDAV  
KIDETVCRITAYPRNTYPTYSPTIRISQYOCGLICARDACSLCIRTPSAAYESR  
IVAGSGSAHPPIWPGDNVINTPAMGLHITPRAPBOCHTHTTDEYSASHTHCYPA  
YSPRTAFIEDEFTDPPDEVEGMDYMAAPVPTPTQETAGGDPVPTQTARDHRP  
PDNLTAFIEDEFTDPPDEVEGMDYMAAPVPTPTQETAGGDPVPTQTARDHRP  
12019..12453  
/gene="OJ1060\_D03.5"  
12019..12453



/gene="OJ1060\_D03.5"  
 /note="hypothetical protein  
 predicted by Glimmer etc."  
 /codon\_start=-1  
 /protein\_id="BAC22264.1"  
 /db\_xref="GI:24414014"  
 /translation="MRCAMINHRLFDHARARVOSAMLLTVNTPSPALAPVNRCH  
 LPNGEKTPKRSRCLHVSHTRAKATPMPIPTPYSSGSGSNARNGRPSRSARAAA  
 AATWGEEDTNGKSDSLPPQPKVPCGACWMPAHAMAP"  
 join(12948..14258,15290..15376,16109..16198)  
 /gene="OJ1060\_D03.6"  
 join(12948..14258,15290..15376,16109..16198)  
 /gene="OJ1060\_D03.6"  
 /note="arm repeat containing protein homolog-like"  
 /codon\_start=-1  
 /protein\_id="BAC22265.1"  
 /db\_xref="GI:24414015"  
 /translation="MVSLSQIPSPQSFCAARSGRAGYSMTTIRSLALQPDSC  
 PGSPHYAAVADAGADSDMNTDVIIDFLSELATAGAPHAAYAKSSANAATE  
 MLEISDFSDSPNSDISELERLAAAAVTPPSDAPVGVADLNEESDLSE  
 AAPLEPPEPVLCVRAICGPDAPAPRAAARIRLAKRSIRRELIGSCAIPALV  
 PLRSTDPVAQESAVTALNLSLENNRSATIRAGIKPLVIALRTGTASAKONACA  
 LLSLSGIEENRATIGACAIPLVALLSAGTGGKDALTTILRLCSARNKERRAVSA  
 GAVVPLIHVGERSGTSEKAMVLAISAGIVGRDAVWAGGIPALVETIEDGPARE  
 REFAVVALLQICSECPNRNALVREGAIPLVALLSOGSARAKHKAETLLGYLREQRO  
 GGCGRVEPVAASLAFIVLCVTFSLFTPSGSAASIMWTYSYT"  
 17638..17869  
 /note="5' LTR"  
 /complement(join(19591..19658,20059..20179))  
 /gene="OJ1060\_D03.7"  
 complement(join(19591..19658,20059..20179))  
 /gene="OJ1060\_D03.7"  
 /note="this category is not included in IRGSP standard.  
 hypothetical ORF  
 predicted by GENSCAN"  
 24046..25584  
 /gene="OJ1060\_D03.8"  
 24046..25584  
 /gene="OJ1060\_D03.8"  
 /note="deletion  
 pseudogene, Tamil transposon protein TNP2"  
 /pseudo  
 complement(join(26731..27006,27068..27145))  
 /gene="OJ1060\_D03.9"  
 complement(join(26731..27006,27068..27145))  
 /gene="OJ1060\_D03.9"  
 /note="hypothetical protein  
 predicted by GENSCAN etc."  
 /codon\_start=-1  
 /protein\_id="BAC22266.1"  
 /db\_xref="GI:24414016"  
 /translation="MAARILCAAAVRRCCAAAASGRCVRYVASCGRIRVAVIGV  
 AAESCAATTAASAATRRPAAVWGERIEVGERGRRRRLSPORCGGGGDDAR  
 EATARYDGVGAGMP"  
 28336..31791  
 /gene="OJ1060\_D03.10"  
 28336..31791  
 /gene="OJ1060\_D03.10"  
 /note="probably inactive due to frameshift(s) in CDS  
 pseudogene, hydroxyproline-rich glycoprotein"  
 /pseudo  
 36755..37150  
 /gene="OJ1060\_D03.11"  
 36755..37150  
 /gene="OJ1060\_D03.11"  
 /note="hypothetical protein  
 predicted by gm etc."  
 /codon\_start=-1  
 /protein\_id="BAC22267.1"  
 /db\_xref="GI:24414017"  
 /translation="MILFPIHQTEAHEDGGRGGOYPPVENSARACGGRGAPHATPT  
 PRNRAGRGORPRRSGLTERRFLRGLKPKRYALLIQRILQVADAPHWCGPSRAATSSO

gene  
 CDS  
 38163..39500  
 /gene="OJ1060\_D03.12"  
 38163..39500  
 /gene="OJ1060\_D03.12"  
 /note="contains ESTs  
 AU183957(R0167), D46798(S11701), AU077610(E20630)  
 similar to Arabidopsis thaliana chromosome5, At5g55530  
 unknown protein"  
 /codon\_start=-1  
 /protein\_id="BAC22268.1"  
 Query Match 87.0% Score 17.4; DB 8; Length 102178;  
 Best Local Similarity 94.7% Pred. No. 1.6e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 TGTCAAAACCGTTGTCTC 19  
 Db 84821 TTTCAAAACCGTTGTCTC 84803  
 RESULT 10  
 AC107604/c  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone CH230-196k4, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*  
 AC107604 223885 bp DNA linear HTG 21-SEP-2002  
 Rattus norvegicus  
 AC107604.4 GI:23266137  
 HTG: HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 223885)  
 Muzny,D,Marle, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,  
 Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,  
 Anyalebech,Y, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,  
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,  
 Bialwalo,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,  
 Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,  
 Cardenas,V, Carter,K, Cavazos,I, Caesar,K, Center,A,  
 Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,  
 Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,  
 Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,  
 Delgado,O, Denson,S, Derramo,C, Ding,Y, Dinh,H, Divya,K,  
 Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Davies,K,  
 Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,  
 Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,  
 Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,  
 Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,M, Guevara,M,  
 Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,  
 Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,R,  
 Hernandez,R, Hines,S, Hladun,S,L, Hodgson,N, Hoguez,M,  
 Hollins,B, Howells,S, Huylk,S, Hume,J, Idlebird,D, Jackson,A,  
 Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolyet,A,  
 Karpethy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,  
 Kows,J, Kraft,C,L, Ledow,H, Levan,J, Lewis,L, Li,Z, Liu,J,  
 Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,Z,  
 Lorensunewa,L, Louisedge,H, Lozado,R,J, Lu,X, Ma,J,  
 Maheshwari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangum,A,  
 Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E,  
 Mahoney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,  
 Mlosavljevic,A, Miner,G, Minga,E, Montemayor,J, Moore,S,  
 Morgan,M, Morris,K, Morris,S, Nuntiasa,M, Murphy,M, Nair,L,  
 Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,  
 Nwaokemele,O, Okwunonu,G, Olarunpusogoon,A, Pal,S, Parks,K,  
 Pasternak,S, Paul,H, Perez,A, Perez,L, Plankoch,C,  
 Plopper,F, Polindexter,A, Popovic,D, Plums,E, Pu,L, Li,  
 Puazo,B, Quiroz,J, Rachlin,E, Reeves,K, Register,M,A, Reish,R,  
 Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,  
 Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ritz,S,J,  
 Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,

Shetty, J., Shivartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taboc, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Journal Title: Direct Submission  
Unpublished  
2 (bases 1 to 223885)  
Worley, K.C.  
Submitted (23-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 223885)  
Rat Genome Sequencing Consortium.  
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT  
On Sep 21, 2002 this sequence version replaced gi:21737824.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: GPUL  
Center clone name: CH230-196K4

Summary Statistics  
Assembly program: Phrap, version 0.990329  
Consensus quality: 212075 bases at least Q40  
Consensus quality: 214464 bases at least Q20  
Estimated insert size: 222767; sum-of-contigs estimation  
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces  
is believed to be correct as given, however the sizes  
of the gaps between them are based on estimates that have  
been provided by the submitter.  
This sequence will be replaced  
by the finished sequence as soon as it is available and  
the accession number will be preserved.

1 223885: contig of 223885 bp in length.

FEATURES  
source  
1..223885  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-196K4"  
1..1816  
misc-feature  
/note="wgs\_end\_extension  
clone\_end:5p6"  
1866..2449  
misc-feature  
/note="clone\_boundary  
clone\_end:5p6"

misc-feature  
end..sequence: BH318237"  
complement(1221124..223882)  
/note="clone\_boundary  
clone\_end:77  
site: Ecoli  
end..sequence: BH318236"  
222550..223885  
/note="wgs\_end\_extension  
clone\_end:77"

misc-feature  
BASE COUNT 56904 a 48879 c 49011 g 60401 t 8690 others  
ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 223885;  
Best Local Similarity 94.7%; Pred. No. 1.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCAAAACCTGTGTCC 20  
|||||  
DB 102327 GTCAAAACCTGTGTCC 102309

RESULT 11  
AC094509/C  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-4E17, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
AC094509  
AC094509.6 GI:30466778  
VERSION  
KEYWORDS  
HTG; HTGS; PHASE2; HTGS; DRAFT; HTGS; ENRICHED.  
SOURCE  
Rattus norvegicus (Norway rat)  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS  
1 (bases 1 to 236021)  
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Arguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Guarnatore, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huix, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louiseged, H., Lozardo, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangun, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwunu, G., Olariunsaogun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruit, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, F., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished

2 (bases 1 to 236021)

Worley, K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236021)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:23265564.

## COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

## ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GAND

Center clone name: CH230-4E17

Summary Statistics

Assembly program: Atlas

Consensus quality: 218135 bases at least Q40

Consensus quality: 221412 bases at least Q30

Consensus quality: 223855 bases at least Q20

Estimated insert size: 229824; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

- \* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).
- \* NOTE: This is a 'working draft' sequence. It currently
- \* consists of 1 contigs. Gaps between the contigs
- \* are represented as runs of N. The order of the pieces
- \* is believed to be correct as given, however the sizes
- \* of the gaps between them are based on estimates that have
- \* provided by the submitter.
- \* This sequence will be replaced
- \* by the finished sequence as soon as it is available and
- \* the accession number will be preserved.

1 236021: contig of 236021 bp in length.

Location/Qualifiers

1. 236021

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-4E17"

misc\_feature

1. 2200

misc\_feature

/note="wgs\_end\_extension  
clone\_end:Sp6"  
5478..6115  
/note="clone\_boundary  
clone\_end:Sp6  
site:ECORI  
end\_sequence:BH305262"  
complement(229315..230090)  
/note="clone\_boundary  
clone\_end:T7  
site:ECORI  
end\_sequence:BH305226"  
232901..236021  
/note="wgs\_end\_extension  
clone\_end:T7"

misc\_feature

BASE COUNT 69961 a 46038 c 45434 g 64133 t 10455 others

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 236021;  
Best Local Similarity 94.7%; Pred. No. 1.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTCAAAACCGTGTGTC 20

Db 74712 GTCAAAACCGTGTGTC 74694

## RESULT 12

AC105688/c

LOCUS

DEFINITION

AC105688

VERSION

KEYWORDS

SOURCE

ORGANISM

AC105688

GI:23264260

HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 236062)

Munzy, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anylebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barstead, M., Benamed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E.,

Cardenas, V., Carter, R., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flaggs, N., Forbes, L., Foster, T., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Georgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, T., Haylak, P., Hayes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladik, S. L., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorenshew, L., Louisege, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, J., Morris, K., Morris, S., Munkitase, M., Murphy, M., Nait, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nankervis, O., Okunou, G., Olariunpasegon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Piazzi, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Savary, G., Scherer, S., Scott, G., Shattman, S., Shen, H.,  
Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Tabor, P., Usmani, K.,  
Valas, R., Vera, Y., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 236062)  
Worley, K.C.

Direct Submission  
Submitted (09-JAN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 236062)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Sep 21, 2002 this sequence version replaced gi:21743898.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the  
sequence may extend beyond the ends of the clone and there may be  
contigs that consist entirely of whole genome shotgun sequence  
reads. Both end sequences and whole genome shotgun sequence only  
contigs will be indicated in the feature table.

----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information -----  
Center project name: GNMD  
Center clone name: CH230-11H4  
----- Summary Statistics -----  
Assembly program: Phrap; version 0.990329  
Consensus quality: 200479 bases at least Q40  
Consensus quality: 204791 bases at least Q30  
Consensus quality: 207411 bases at least Q20  
Estimated insert size: 251580; sum-of-contigs estimation  
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drift\_data.html)  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 234224: contig of 234224 bp in length  
\* 234225 234324: gap of unknown length  
\* 234325 236062: contig of 1738 bp in length.

Location/Qualifiers  
1. 236062  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"

FEATURES  
Source

/db\_xref="taxon:10116"  
/clone="CH230-11H4"  
1064..1913  
/note="clone\_boundary  
clone\_end:f7  
site:EcoRI  
end\_sequence:BH341597"  
179267..180672  
/note="wgs\_contig"  
complement(231465..232270)  
/note="clone\_boundary  
clone\_end:sp6  
site:EcoRI  
end\_sequence:BH341599"

BASE COUNT 65632 a 42528 c 41770 g 59963 t 26169 others  
ORIGIN

Query Match 87.0% Score 17.4; DB 2: Length 236062;  
Best Local Similarity 94.7% Pred. No. 1.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCAAAACGCTGTGCTCC 20  
||||||| |||||||||  
LOCUS AJ414153 258050 bp DNA linear BCT 21-NOV-2002  
DEFINITION Yersinia pestis strain CO92 complete genome; segment 13/20.  
ACCESSION AJ414153 AL590842  
VERSION AJ414153.1 GI:15980576  
KEYWORDS  
SOURCE  
ORGANISM  
Yersinia pestis CO92  
Yersinia pestis CO92  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Yersinia.  
1 (bases 1 to 258050)  
Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T.G.,  
Prentice, M.B., Sebahia, M., James, K.D., Churcher, C., Mungall, K.L.,  
Baker, S., Basham, D., Bentley, S.D., Brooks, K., Cerdano-Tarraga, A.M.,  
Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G.,  
Fellwell, T., Hamlin, N., Holroyd, S., Jagels, K., Leather, S.,  
Karlshew, A.V., Moule, S., Oyston, P.C.F., Quail, M., Rutherford, K.,  
Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G.  
Genome sequence of Yersinia pestis, the causative agent of plague  
Nature 413 (6855), 523-527 (2001)  
21470413  
MEDLINE 11586360  
PUBMED 2 (bases 1 to 258050)  
REFERENCE Parkhill, J.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (04-OCT-2001) Submitted on behalf of the Yersinia  
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,  
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk

COMMENT  
Notes:  
Details of Y. pestis sequencing at the Sanger Centre are available  
on the World Wide Web.  
(URL, http://www.sanger.ac.uk/Projects/Y\_pestis/).

Location/Qualifiers  
1. 258050  
/organism="Yersinia pestis CO92"  
/mol\_type="genomic DNA"  
/strain="CO92"  
/db\_xref="taxon:214092"  
/note="biovar: Orientalis"  
396..590  
/gene="YPO2590"  
396..590  
/gene="YPO2590"  
/note="no significant database hits"  
/codon\_start=1  
/transl\_table=11

FEATURES  
Source  
gene  
CDS

gene	/product="hypothetical protein" /protein_id="CAC92833.1" /db_xref="GI:15980577" /translation="MSTLEKLYSAVCHTSLDPVASTVAFMSEQKNIDVDKIEATLSV DERHFFRRRTLYVRDILYRPL" complement(608..790) /gene="YPO2591" /note="no significant database hits" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAC92834.1" /db_xref="GI:15980578" /translation="MKTSHRAISTYREARKAIRSGYDKRIPKGIKTGFKDKPPDEM TTIEVHTIOVALILSV" complement(972..1937) /gene="YPO2592" complement(972..1937) /note="Similar to Streptococcus pneumoniae transmembrane protein CAP35W TR:086896 (EMBL:AJ006986) (332 aa) fasta scores: E(): 9.7e-14, 25.4% id in 335 aa, and to internal region of Campylobacter jejuni probable enterochelin uptake permease CeuC TR:09PM06 (EMBL:AL139078) (312 aa) fasta scores: E(): 0.13, 23.3% id in 296 aa" /codon_start=1 /transl_table=11 /product="putative membrane protein" /protein_id="CAC92835.1" /db_xref="GI:15980579" /translation="MRNYGLNKLKTLSCPSANTFYSAKCTEQGLSGEWGVLYFL STIANPLRFMIIGYIDSDNDKIDKIDFKRLKITTITIFWNVLFYFINDGKRRYFL OSWLFSTAIYILNPIISLIKLSNRTAIIITLSCVIFSVTIDLISAREPRLIDFP OYRFMTAFYFMTGRFLCSRICQMTKLPKYRLAKLILLPTSLISMYEESFMISY YKTVNASEYFLDHFVILSLCLFVIFDNDTQYEWIKTKLTIISPMIGVYILHDGIF YFIASVNVSDVTLRFLLISYFASVLLSRALLNKTSRISF" complement(1005..1070) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1101..1166) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1206..1262) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1308..1361) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1401..1451) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1497..1547) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1569..1634) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" complement(1665..1715) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0"
misc_feature	complement(1755..1820) /gene="YPO2592" /note="One of 9 probable transmembrane helices predicted for YPO2592 by TMHMM2.0" 2360..2977 /gene="YPO2593" /gene="YPO2593" /note="Similar to Salmonella typhimurium fimbriae W protein FimW SM:FIMW_SALTY (P37928) (198 aa) fasta scores: E(): 0.0066, 21.9% id in 178 aa, and to C-terminus of Erwinia amylovora activator of exopolysaccharide synthesis, RcsB TR:P96320 (EMBL:Y09848) (215 aa) fasta scores: E(): 0.0003, 43.7% id in 71 aa" /codon_start=1 /transl_table=11 /product="putative LuxR-family regulatory proteins" /protein_id="CAC92836.1" /db_xref="GI:15980580" /db_xref="SPTRMBL:Q8ZDH5" /translation="MGIAKMSCMIVDNDKQYQTLGAVSIYKRIPTSLGFKKEIFYKRR AFYSADVIFGVDESPFEALRDLDPKPAEDVFLICDARNSLPGIIPRSDVTMIF REDLSLEYNSKIAIKYKRRKPGFDNLIERATVILHYSERQYLPNNIVYALKLN EGFSGGDIAKILKSEKTSVQSKSAKKLIGARDVELIKFMFK" 2792..2968 /gene="YPO2593" /note="Pfam match to entry PF00196 GERE, Bacterial regulatory proteins, LuxR family, score 21.90, E-value 0.00026" 2846..2911 /gene="YPO2593" /note="Predicted helix-turn-helix motif with score 1182 (+3.21 SD) at aa 163-184, sequence FSGDIAKILKSEKTSVQSKR" complement(3297..3668) /gene="YPO2594" complement(3297..3668) /gene="YPO2594" /note="no significant database hits" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAC92837.1" /db_xref="GI:15980581" /db_xref="SPTRMBL:Q8ZDH4" /translation="MIKGAIVTEPEPTNTGLSAFVAKVAIEKRYIFIANERIKLTS IKSFNTFKVEVPAQAEVDKANTILSVNLNKKIRPQEDISEKTIPLKMLITTFQNVV YTPSEHVDIDISYAKRKS" 4054..4263 /gene="cspE" /note="synonym: YPO2595" 4054..4263 /gene="cspE" /note="Similar to Escherichia coli probable cold shock protein CSPB SW:CSPB_ECOLI (P36997) (68 aa) fasta scores: E(): 1e-25, 94.1% id in 68 aa, and to Vibrio cholerae hypothetical protein TR:AACT2388 (EMBL:AF098260) (69 aa) fasta scores: E(): 3.1e-26, 94.2% id in 69 aa. Similar to YPO1746 (84.1% identity in 69 aa overlap), YPO3644 (73.6% identity in 72 aa overlap), YPO3643 (72.2% identity in 72 aa overlap), YPO2659 (63.9% identity in 72 aa overlap) and YPO1398 (53.6% identity in 69 aa overlap)." /codon_start=1 /transl_table=11 /product="putative cold shock protein" /protein_id="CAC92838.1" /db_xref="GI:15980582" /db_xref="SPTRMBL:Q8ZDH3" /translation="MSKIKSVKWFNENSGKGFITPDGSKNDVYVHPSATASNGFKTL AEGQVREFEITNGAKGPSANAVAI" 4060..4260 /gene="cspE" /note="Pfam match to entry PF00313 GSD, 'Cold-shock'

misc\_feature DNA-binding domain, score 152.30, E-value 8.6e-42"  
4102..4161  
/gene="cspe"  
/note="PS00352 'Cold-shock' DNA-binding domain signature."  
gene complement(4686..5069)  
/gene="crcb"  
/note="synonym: YPO2596"  
CDS complement(4686..5069)  
/gene="crcb"

Query Match 87.0%; Score 17.4; DB 1; Length 258050;  
Best Local Similarity 94.7%; Pred. No. 1.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTTGTGTC 19  
|||||  
DB 194210 TGTCAAAAACCGTTATGTC 194228

RESULT 14  
E05251 20 bp DNA linear PAT 29-SEP-1997  
LOCUS Part of DNA sequence of Human Papillomavirus 33.  
ACCESSION E05251  
VERSION E05251.1 GI:2173441  
KEYWORDS JP 1993192200-A/21.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.  
TITLE DETECTION OF HUMAN PAPILLOMA VIRUS  
JOURNAL Patent: JP 1993192200-A 21 03-AUG-1993;  
TAKARA SHUZO CO LTD  
COMMENT OC Artificial gene  
OS Human papillomavirus  
PN JP 1993192200-A/21  
PD 03-AUG-1993  
PF 19-AUG-1991 JP 1991230839  
PR 20-AUG-1990 JP 90P 217067  
PI OKAZAWA KAZUHIRO, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI  
FUKUSHIMA MICHIO,  
FUJINAGA KET,  
PC C12Q1/70,C12Q1/68;  
CC strandedness: Single;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No.  
FEATURES Location/Qualifiers  
source 1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 4 a 5 c 4 g 7 t  
ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 20;  
Best Local Similarity 90.0%; Pred. No. 3.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTTGTGTC 20  
|||||  
DB 1 TGTCAAAAGACCTTGTGTC 20

RESULT 15  
AF234529 450 bp DNA linear VRL 28-MAR-2000  
LOCUS Human papillomavirus type 58 strain E6HK1 E6 protein gene, complete  
DEFINITION cds.  
ACCESSION AF234529  
VERSION AF234529.1 GI:7331167

KEYWORDS Human papillomavirus type 58  
SOURCE Human papillomavirus type 58  
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
REFERENCE 1 (bases 1 to 450)  
AUTHORS Chan,P.K.S., Lam,C.W., Li,W.H., Chan,M.Y.M., Cheung,J.L.K. and  
Cheng,A.F.  
TITLE Nucleotide sequence variation of the E6 gene of human  
papillomavirus type 58 from Chinese women with normal cervixes and  
cancerous lesions  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 450)  
AUTHORS Chan,P.K.S., Lam,C.W., Li,W.H., Chan,M.Y.M., Cheung,J.L.K. and  
Cheng,A.F.  
TITLE Direct Submission  
JOURNAL Submitted (15-FEB-2000) Department of Microbiology, The Chinese  
University of Hong Kong, Prince of Wales Hospital, Shatin, NT.,  
Hong Kong  
FEATURES Location/Qualifiers  
source 1..450  
/organism="Human papillomavirus type 58"  
/mol\_type="genomic DNA"  
/strain="E6HK1"  
/isolate="QEH61"  
/db\_xref="taxon:10598"  
/country="China"  
/note="from cervical scrape of patient with invasive  
squamous cell carcinoma of the cervix"  
CDS 1..450  
/note="oncoprotein; transforming; trans-activating;  
p53-binding"  
/codon\_start=1  
/product="E6 protein"  
/protein\_id="AAF60305.1"  
/db\_xref="GI:7331168"  
/translation="MFQDAEKPRLHDLCALETSVHEIEKCECKKTKTQRESYVD  
FVFDLRIVRDGNFPAVCVKCLRLISSEYRHVNSLYGDTLEQTLKCLNETLIR  
CIICORPLCPQEKRRHVDLNRFPNHSISGMTGRCACVWRPRRQTV"  
BASE COUNT 158 a 60 c 104 g 128 t  
ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 450;  
Best Local Similarity 90.0%; Pred. No. 3.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTTGTGTC 20  
|||||  
DB 316 TGTCAAAAGACCTTGTGTC 335

Search completed: August 23, 2003, 11:11:33  
Job time : 995 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 05:06:07 ; Search time 206.5 Seconds  
(without alignments)  
261.447 Million cell updates/sec

Title: US-09-927-585A-1  
Perfect score: 20  
Sequence: 1 tgcataaacctgtgtcc 20

Scoring table: GAPOP 10.0, Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*  
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	14	AAQ48566	HPV E6/7 region pr
2	20	100.0	20	24	AA18612	Middle-high hazard
3	20	100.0	20	24	AB75144	Oncogenic HPV regi
4	18.4	92.0	20	14	AAQ48556	HPV E6/7 region pr
5	18.4	92.0	1023	14	AAQ48575	HPV E6/7 Genomic r
6	16.8	84.0	20	14	AAQ48560	HPV E6/7 region pr
7	16.8	84.0	540	24	AAQ1435	Yeast selected int
8	16.8	84.0	541	24	AB562861	Selected Interacti

C	9	16.8	84.0	766	22	AAH70910	Human cervical can
C	10	16.8	84.0	774	10	AAQ91602	Partial nucleotide
C	11	16.8	84.0	3443	24	AAQ76303	S. cerevisiae BAX-
C	12	16	80.0	530	22	AAE25493	Nucleotide sequenc
C	13	16	80.0	1447	22	AAE25498	Nucleotide sequenc
C	14	16	80.0	265118	22	AAH41227	Pyrococcus abyssi
C	15	15.8	79.0	611	24	ABO55587	Human ovarian anti
C	16	15.8	79.0	1409	25	ABT19257	Aspergillus fumiga
C	17	15.8	79.0	1518	25	ABT21077	Aspergillus fumiga
C	18	15.8	79.0	1521	25	ABT20479	Aspergillus fumiga
C	19	15.8	79.0	1521	25	ABT18663	Aspergillus fumiga
C	20	15.8	79.0	3518	25	ABT18069	Aspergillus fumiga
C	21	15.8	79.0	3520	25	ABT18069	Aspergillus fumiga
C	22	15.8	79.0	29046	23	ABH10478	Drosophila melanog
C	23	15.8	79.0	349980	22	AAH41225	Pyrococcus abyssi
C	24	15.4	77.0	3339	22	AAH41225	DNA encoding novel
C	25	15.4	77.0	13309	23	AAK83980	Human immune/hema
C	26	15.2	76.0	20	14	AAQ48549	HPV E6/7 region pr
C	27	15.2	76.0	29	16	AAQ76065	HPV primer set 13,
C	28	15.2	76.0	100	19	AAQ39991	E6 oncogene from H
C	29	15.2	76.0	138	24	ABO73531	HPV-PTM related ol
C	30	15.2	76.0	456	19	ABA27673	Human papillomavir
C	31	15.2	76.0	477	22	ABA58702	Human foetal liver
C	32	15.2	76.0	477	22	AAK32556	Probe #6139 for ge
C	33	15.2	76.0	477	22	AAK32556	Human brain expres
C	34	15.2	76.0	477	22	AAK32556	Human bone marrow
C	35	15.2	76.0	477	22	AAK32556	Probe #6055 for ge
C	36	15.2	76.0	477	22	AAI18381	Human liver single
C	37	15.2	76.0	477	23	AB532267	Human genome-deliv
C	38	15.2	76.0	477	24	AB507346	HPV16 E6 siRNA der
C	39	15.2	76.0	477	25	AAI54430	Human papilloma vi
C	40	15.2	76.0	477	25	AAI54445	Mammalian vestibul
C	41	15.2	76.0	519	17	AAI18384	HPV16 E6/E7 encodi
C	42	15.2	76.0	519	17	AAI18384	Drosophila melanog
C	43	15.2	76.0	525	16	AAQ5470	
C	44	15.2	76.0	570	16	AAQ5470	
C	45	15.2	76.0	642	23	ABL04975	

## ALIGNMENTS

RESULT 1	AAQ48566	standard; DNA; 20 BP.
ID	AAQ48566	
AC	AAQ48566	
XX	22-FEB-1994 (first entry)	
XX	HPV E6/7 region probe.	
DE	Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.	
XX	Synthetic.	
OS	JP05192200-A.	
PN	03-AUG-1993.	
XX	19-AUG-1991; 91JP-0230839.	
XX	20-AUG-1990; 90JP-0217067.	
XX	(TAKI ) TAKARA SHUZO CO LTD.	
XX	WPI: 1993-277497/35.	
DR	Detecting benign and/or malignant human papilloma virus - by	
XX	detecting DNA sequence of E6 and/or E7 region of human papilloma	
PT	virus	
XX	Disclosure; Page 15; 18pp; Japanese.	

XX The probe is used to detect benign and/or malignant human papilloma  
CC virus. The probe binds to the E6 and/or E7 region of the virus.  
XX  
Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 TGTCAAAACCGTGTGTCC 20  
1 TGTCAAAACCGTGTGTCC 20

RESULT 2  
AAS18612  
ID AAS18612 standard; DNA; 20 BP.  
XX  
AC AAS18612;

XX 26-FEB-2002 (first entry)

XX Middle-high hazard human papillomavirus (HPV) E6, PCR primer.

XX Human Papillomavirus; HPV, early gene; E6; PCR primer;  
KW middle-high hazard; ss.

XX Human papillomavirus.

XX W020018594-A2.

XX 15-NOV-2001.

XX 02-MAY-2001; 2001MO-IB00771.

XX 05-MAY-2000; 2000IT-MC00091.

XX (BIOA-) BIOANALIST CENT SUD DI PERSEU SINIBLADO.

XX Perseu S, De Montis A, Floris NM;

XX WPI; 2002-075253/10.

XX Identifying viral DNA of human papilloma virus in cellular material  
PT collected from an individual, by using polymerase and identifying  
PT sequences complementary to regions E of the viral genome using a  
PT reagent.

XX Claim 9; Page 12; 17pp; English.

XX The invention describes a novel method of identifying viral DNA of human  
CC papillomavirus (HPV) in cellular material collected from an individual.  
CC The method comprises extracting DNA from the cellular material,  
CC amplifying the DNA by using a polymerase, and identifying sequences  
CC complementary to sequences in the regions E (early genes) e.g. E6 and E7  
CC of the HPV genome using an identifying reagent. The presence of HPV in  
CC cells is now based on identifying a reagent. The presence of HPV in  
CC region of the genome. Identification of E6, E7 and I1 make it possible to  
CC distinguish the presence of HPV virus in every stage of infection and to  
CC distinguish between virus strains having low hazard and virus strains  
CC having middle/high hazard. This is the primer used to identify the E6  
CC region of middle-high hazard HPV strains in infected cells, described in  
CC the method of the invention.

XX Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 TGTCAAAACCGTGTGTCC 20  
XXXXXXXXXXXXXXXXXXXX

DB 1 TGTCAAAACCGTGTGTCC 20

RESULT 3  
ABZ75144

ID ABZ75144 standard; DNA; 20 BP.

XX  
AC ABZ75144;

XX 30-MAY-2003 (first entry)

XX Oncogenic HPV region E6-E7 PCR primer VI-3, SEQ ID NO:18.

XX Reaction mixture; desiccated; dried; enzymatic; stabilised;  
KW condensation reaction inhibitor; polymer mesh; nucleic acid manipulation;  
KW amplification; sequencing; hybridisation; restriction analysis;  
KW diagnosis; ready-to-use; hot start reaction; HPV; identification;  
XX oncogenic type; E6-E7; PCR; primer; ss.

XX Human papillomavirus type 16.

XX Human papillomavirus type 18.

XX Human papillomavirus type 31.

XX Human papillomavirus type 33.

XX Human papillomavirus type 35.

XX Human papillomavirus type 32b.

XX Human papillomavirus type 58.

XX W0200272002-A2.

XX 19-SEP-2002.

XX 11-MAR-2002; 2002MO-ES00109.

XX 12-MAR-2001; 2001ES-0000569.

XX (BIOT-) BIOTOLS BIOTECHNOLOGICAL & MEDICAL LAB.

XX Franco De Sarabia Rosado PM, Lmones Lopez G, Madejon Seiz A;

XX Marn Alberdi MD;

XX WPI; 2003-067358/06.

XX Stabilized reaction mixture containing an enzyme, useful for performing  
PT nucleic acid reactions, includes three-component stabilizing mixture  
PT and is at least partially dried -  
XX

Example VI; Page 41; 76pp; Spanish.

XX The invention relates to a method for preparing a stabilised reaction  
CC mixture which contains at least one enzyme and which is at least  
CC partially dried. The method involves removing the water from an aqueous  
CC reaction mixture containing an enzyme and a three-component stabiliser  
CC solution until the residual moisture content is 30% or less. The  
CC three-component stabiliser solution used in the method comprises an  
CC agent which protects against desiccation, an inhibitor of condensation  
CC reactions between carbonyl or carboxy groups and amino or phosphate  
CC groups, and an inert polymer that forms a mesh structure that inhibits  
CC the mobility of the dried reactants. The dried reaction mixtures can be  
CC used as "ready-to-use" mixtures for performing a wide range of nucleic  
CC acid manipulations such as amplification, sequencing, hybridisation  
CC and/or restriction analysis, e.g., for the diagnostic detection of  
CC pathogens or mutations. The dried reaction mixtures, already deposited in  
CC components needed to perform a particular reaction, already deposited in  
CC a reaction vessel, which eliminates the need for multiple additions  
CC (which are sources of errors and contamination), thereby improving  
CC repeatability and reliability. The dried reaction mixtures can be  
CC transported and stored at ambient temperature without significant loss  
CC of activity, and are suitable for "hot start" reactions. Sequences  
CC ABZ75142-ABZ75145 represent PCR primers used in an exemplification to  
CC amplify regions of human papillomavirus (HPV) genomes in order to  
CC characterise their type. Primers ABZ75142-ABZ75143 are PCR primers used  
CC to amplify a 450 bp region designated I1 which is shared by all HPV  
CC types, while PCR primers ABZ75144-ABZ75145 are used to amplify a 250 bp



CC region designated E6-E7 from oncogenic HPV types only.  
 XX Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 25; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTGTGTC 20  
 |||||||  
 DB 1 TGTCAAAAACCGTGTGTC 20

RESULT 4  
 AAQ48556  
 ID AAQ48556 standard; DNA; 20 BP.

AC AAQ48556;

DT 22-FEB-1994 (first entry)

DE HPV E6/7 region probe.

KW Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.

OS Synthetic.

PN JP05192200-A.

PD 03-AUG-1993.

PF 19-AUG-1991; 91JP-0230839.

PR 20-AUG-1990; 90JP-0217067.

PA (TAKI ) TAKARA SHUZO CO LTD.

DR WPI; 1993-277497/35.

PT Detecting benign and/or malignant human papilloma virus - by

PT detecting DNA sequence of E6 and/or E7 region of human papilloma

PS Disclosure; Page 14; 18pp; Japanese.

CC The probe is used to detect benign and/or malignant human papilloma

CC virus. The probe binds to the E6 and/or E7 region of the virus.

XX Sequence 20 BP; 4 A; 5 C; 5 G; 6 T; 0 other;

Query Match 92.0%; Score 18.4; DB 14; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTGTGTC 20  
 |||||||  
 DB 1 TGTCAAAAACCGTGTGTC 20

RESULT 5  
 AAQ48575  
 ID AAQ48575 standard; DNA; 1023 BP.

AC AAQ48575;

DT 22-FEB-1994 (first entry)

DE HPV E6/7 Genomic region.

KW Human papilloma virus; HPV; benign; malignant; E6; E7; ss.

OS Synthetic.

XX

EH Key Location/Qualifiers  
 FT CDS 250..696

FT /\*tag= a

FT /label= ORF-1

FT /note= "E6 region"

FT 702..998

FT /\*tag= b

FT /label= ORF-2

FT /note= "E7 region"

PN JP05192200-A.

PD 03-AUG-1993.

PF 19-AUG-1991; 91JP-0230839.

PR 20-AUG-1990; 90JP-0217067.

PA (TAKI ) TAKARA SHUZO CO LTD.

DR WPI; 1993-277497/35.

DR P-PSDB; AAR40919, AAR48202.

PT Detecting benign and/or malignant human papilloma virus - by

PT detecting the DNA sequence of E6 and/or E7 region of human papilloma

PS Claim 1; Page 16-17; 18pp; Japanese.

CC Probes were designed to detect benign and/or malignant human papilloma

CC virus. The probes bind to the E6 and/or E7 region of human papilloma

XX Sequence 1023 BP; 355 A; 170 C; 227 G; 271 T; 0 other;

Query Match 92.0%; Score 18.4; DB 14; Length 1023;  
 Best Local Similarity 95.0%; Pred. No. 18;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTGTGTC 20  
 |||||||  
 DB 565 TGTCAAAAACCGTGTGTC 584

RESULT 6  
 AAQ48560  
 ID AAQ48560 standard; DNA; 20 BP.

AC AAQ48560;

DT 22-FEB-1994 (first entry)

DE HPV E6/7 region probe.

KW Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.

OS Synthetic.

PN JP05192200-A.

PD 03-AUG-1993.

PF 19-AUG-1991; 91JP-0230839.

PR 20-AUG-1990; 90JP-0217067.

PA (TAKI ) TAKARA SHUZO CO LTD.

DR WPI; 1993-277497/35.

PT Detecting benign and/or malignant human papilloma virus - by

PT detecting DNA sequence of E6 and/or E7 region of human papilloma

XX virus



PN WO200142467-A2.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 08-DEC-2000; 2000MO-US33312.  
 XX  
 PR 08-DEC-1999; 99US-0169681.  
 PR 21-DEC-1999; 99US-0171350.  
 PR 14-MAR-2000; 2000US-0189315.  
 PR 12-MAY-2000; 2000US-0203791.  
 PR 09-JUN-2000; 2000US-0210600.  
 PR 21-JUL-2000; 2000US-0220114.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Deeds J, Berger A, Zhao X;  
 XX  
 DR WPI; 2001-375006/39.  
 XX  
 PT New isolated nucleic acid for diagnosing and treating cervical cancer  
 PT and for assessing and detecting compounds for treating the cancer -  
 XX  
 PS Claim 1; Page 462; 1051pp; English.  
 XX  
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
 CC polypeptides are useful: to assess if a patient is afflicted with  
 CC cervical cancer or has a pre-malignant condition; to monitor the  
 CC progression of cervical cancer or a premalignant condition in a patient;  
 CC and to select and/or assess the efficacy of a compound or therapy for  
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
 CC useful for gene therapy.  
 XX  
 SQ Sequence 766 BP; 197 A; 158 C; 159 G; 244 T; 8 other;  
 XX  
 QY Query Match 84.0%; Score 16.8; DB 22; Length 766;  
 XX Best Local Similarity 90.0%; Pred. No. 1.1e+02;  
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 DB 1 TGTCAAAAACCGTTGTGTC 20  
 734 TGTCAAAAGCCCTTGTC 715  
 XX  
 RESULT 10  
 AAN91602  
 ID AAN91602 standard; DNA; 774 BP.  
 XX  
 AC AAN91602;  
 XX  
 DT 17-JUL-1990 (first entry)  
 XX  
 DE Partial nucleotide sequence (5' end) of human papilloma virus (HPV)  
 DE type 33 (HPV-33).  
 XX  
 KW Human papilloma virus; type 33; in situ hybridisation assay;  
 KW cellular smear; cervical carcinoma.  
 XX  
 OS Human papilloma virus.  
 XX  
 PN WO8902934-A.  
 PD 06-APR-1989.  
 XX  
 PF 30-SEP-1988; 88WO-US03367.  
 XX  
 PR 02-OCT-1987; 87US-0103979.  
 XX  
 PA (MICR-) MICROPROBE CORP.  
 XX  
 PI Schwartz DE, Adams TH;  
 XX  
 DR WPI; 1989-114406/15.

XX  
 PT Hybridisation test for human papilloma virus in cell smears -  
 PT by reaction with long labelled probe specific for particular  
 PT virus types, esp. for examining cervical smears  
 XX  
 PS Disclosure; 39pp; English.  
 XX  
 CC The patent is for a rapid in situ hybridisation assay for detecting and  
 CC typing human papilloma virus (HPV) in non-frozen cellular smears fixed to  
 CC a support in absence of aldehyde-based crosslinking reagents. The assay  
 CC comprises: (1) combining nucleic acid in the sample with at least one  
 CC detectable probe able to hybridise with 1 or more HPV types; and (2)  
 CC detecting presence or absence of hybrid complexes. Opt. several probes  
 CC are used, eg one for HPV types 6 and 11, associated with benign warts,  
 CC and one for types 16, 18, 31, 33 and 35, associated with cervical  
 CC cancer. The assay can differentiate between HPV types. It is esp. used  
 CC as a secondary test. The probes can be synthesised or cloned.  
 XX  
 SQ Sequence 774 BP; 270 A; 132 C; 163 G; 209 T; 0 other;  
 XX  
 QY Query Match 84.0%; Score 16.8; DB 10; Length 774;  
 XX Best Local Similarity 90.0%; Pred. No. 1.1e+02;  
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 DB 1 TGTCAAAAACCGTTGTGTC 20  
 424 TGTCAAAAGACCTTTGTGTC 443  
 XX  
 RESULT 11  
 ABQ76303/C  
 ID ABQ76303 standard; cDNA; 3443 BP.  
 XX  
 AC ABQ76303;  
 XX  
 DT 21-NOV-2002 (first entry)  
 XX  
 DE S. cerevisiae BAX-associated cDNA fragment spq ID 31.  
 XX  
 KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virocidic;  
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;  
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischemia;  
 KW neurodegeneration; cell death; ss.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO200264766-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 21-DEC-2001; 2001MO-EP15398.  
 XX  
 PR 22-DEC-2000; 2000EP-0870318.  
 PR 04-JAN-2001; 2001EP-0870002.  
 PR 09-JAN-2001; 2001EP-0870003.  
 XX  
 PA (JANC ) JANSSEN PHARM NV.  
 XX  
 PI Contreras RH, Eberhardt I, Luyten WHM, Reekmans RJ;  
 XX  
 DR WPI; 2002-667002/71.  
 DR P-PSDB; ABG93037.  
 XX  
 PT New isolated nucleic acid representing a synthetic Bax gene, useful as  
 PT medicament for treating, preventing and/or alleviating yeast or fungal  
 PT infections or proliferative disorders, or for preventing apoptosis in  
 PT certain diseases  
 XX  
 PS Claim 36; Figure 1; 344pp; English.  
 XX  
 CC This invention describes a novel nucleic acid representing a synthetic  
 CC Bax gene. The Bax gene of the invention is useful for identifying  
 CC Bax-resistant yeast or fungi, identifying, or obtaining and identifying

CC Candida spp. sequences that are differentially expressed in a pathway  
CC eventually leading to programmed cell death or identifying inhibitors or  
CC inhibitor sequences of Bax-induced cell death. The products of the  
CC invention have cytostatic, fungicide, immunosuppressive, virucide and  
CC vasotropic activity and can be used in vaccines or for gene therapy. The  
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,  
CC antisense molecules and antibodies are useful as medicaments or in  
CC preparing a medicament for treating, preventing and/or alleviating  
CC diseases associated with yeast or fungi or proliferative disorders, such  
CC as cancer, or for preventing apoptosis in certain diseases. The compounds  
CC or polypeptides, or the genetically modified organism are useful for  
CC preparing a medicament for modifying the endogenic flora of humans and  
CC other mammals. The vaccine is useful for immunising against yeast or  
CC fungal infections. Apoptosis-related diseases include autoimmune disease,  
CC ischaemia, diseases related with viral infections or neurodegenerations.  
CC This sequence represents a polynucleotide associated with the Bax gene  
CC described in the disclosure of the invention.

XX  
SQ Sequence 3443 BP; 1229 A; 718 C; 650 G; 846 T; 0 other;

Query Match 84.0%; Score 16.8; DB 24; Length 3443;  
Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTTGTCTC 20  
||| ||||| ||||| |||  
DB 2900 TGTGAAAACCGTTGTCTC 2881

RESULT 12  
AAF25493/C  
ID AAF25493 standard; DNA; 530 BP.  
XX  
AC AAF25493;  
XX  
DT 15-MAY-2001 (first entry)  
XX  
DE Nucleotide sequence of a rice AMP deaminase enzyme.  
XX  
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;  
XX inosine; ss.  
XX  
OS Oryza sativa.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 202..312  
FT /\*tag= a  
FT /product= "AMP deaminase"  
XX  
PN WO200109305-A2.  
XX  
PD 08-FEB-2001.  
XX  
PF 28-JUL-2000; 2000WO-US21009.  
XX  
PR 30-JUL-1999; 99US-0146473.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;  
XX  
DR WPI; 2001-159866/16.  
XX P-PSDB; AAB31953.  
XX  
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used  
XX e.g. for designing or identifying herbicides that inhibit the enzyme  
XX activities, and as probes for genetic or physical mapping -  
XX  
PS Claim 2; Page 53; 72pp; English.  
XX  
CC The present sequence encodes an AMP deaminase. The specification also  
CC describes adenosine deaminase. These enzymes convert adenosine to

CC inosine. Mutations in these genes cause disruptions in then salvage and  
CC catabolism of adenosine and AMP. In humans, this may lead to death of  
CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for  
CC genetically and physically mapping genes that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).

XX  
SQ Sequence 530 BP; 156 A; 123 C; 98 G; 150 T; 3 other;

Query Match 80.0%; Score 16; DB 22; Length 530;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTTGT 16  
||| ||||| ||||| |||  
DB 50 TGTCAAAACCGTTGT 35

RESULT 13  
AAF25498/C  
ID AAF25498 standard; DNA; 1447 BP.  
XX  
AC AAF25498;  
XX  
DT 15-MAY-2001 (first entry)  
XX  
DE Nucleotide sequence of a soybean adenosine deaminase enzyme.  
XX  
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;  
XX inosine; ss.  
XX  
OS Glycine max.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 188..1255  
FT /\*tag= a  
FT /product= "adenosine deaminase"  
XX  
PN WO200109305-A2.  
XX  
PD 08-FEB-2001.  
XX  
PF 28-JUL-2000; 2000WO-US21009.  
XX  
PR 30-JUL-1999; 99US-0146473.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;  
XX  
DR WPI; 2001-159866/16.  
XX P-PSDB; AAB31958.  
XX  
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used  
XX e.g. for designing or identifying herbicides that inhibit the enzyme  
XX activities, and as probes for genetic or physical mapping -  
XX  
PS Claim 2; Page 67-68; 72pp; English.  
XX  
CC The present sequence encodes an adenosine deaminase. The specification  
CC also describes adenosine deaminase. These enzymes convert adenosine to  
CC inosine. Mutations in these genes cause disruptions in then salvage and

CC catabolism of adenosine and AMP. In humans, this may lead to death of  
CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for  
CC genetically and physically mapping genes, that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).

SO Sequence 1447 BP; 442 A; 283 C; 324 G; 398 T; 0 other;

Query Match Best Local Similarity 80.0%; Score 16; DB 22; Length 1447;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGT 16  
|||||  
DB 57 TGTCAAAACCGTGT 42

RESULT 14

AAH41227/c  
ID AAH41227 standard; DNA; 265118 BP.

AC AAH41227;

DT 29-OCT-2001 (first entry)

DE Pyrococcus abyssi genomic fragment #6.

KM Hyperthermophilic archaeon; hyperthermophilic protein; ds.

OS Pyrococcus abyssi.

FN Key Location/Qualifiers

FT misc-feature 1..49980

FT /tag- a /note- "This sequence overlaps with the 3' end of

FT AAH41226"

PN FR2792651-A1.

PD 27-OCT-2000.

PF 21-APR-1999; 99FR-0005034.

PR 21-APR-1999; 99FR-0005034.

PA (CNRS ) CNRS CENT NAT RECH SCI.

PI (IFRE-) IFREMER INST FR RECH EXPL MER.

PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

PI Queirelrou J, Weissenbach J, Saurin W, Hellig R;

DR WPI; 2001-126236/14.

PT New nucleotide sequences isolated from Pyrococcus abyssi encode

PS proteins useful in industry -

XX Claim 1; Page 593-665; 1657pp; French.

CC The present invention relates to the genomic sequence of Pyrococcus

CC abyssi and P. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a

CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal

CC vents. The present sequence is a fragment of the genomic sequence of P.

CC abyssi. The 5' end of this sequence overlaps with the 3' end of AAH41226.

CC The proteins of the present invention have various potential industrial  
CC uses, since the proteins are stable at very high temperatures, some up to  
CC 110 degrees centigrade.  
CC Note: This patent is in the same patent family as WO200065062, which  
CC contains additional sequences as shown in AAB99132-AAB99143,  
CC AAH75903-AAH75920 and AAG66436.

SO Sequence 265118 BP; 75226 A; 61302 C; 54005 G; 74585 T; 0 other;

Query Match Best Local Similarity 80.0%; Score 16; DB 22; Length 265118;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGT 16  
|||||  
DB 231084 TGTCAAAACCGTGT 231069

RESULT 15

ABQ55587/c  
ID ABQ55587 standard; cDNA; 611 BP.

AC ABQ55587;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOFV22 cDNA, SEQ ID NO:1467.

KM Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KM ovarian cancer; breast cancer; tumour; reproductive system disorder;

KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KM PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;

KM inflammatory condition; immune disorder; blood disorder;

KM cardiovascular disorder; respiratory disorder; neurological disorder;

KM gastrointestinal disorder; urinary system disorder; drug screening;

KM gene therapy; chromosome mapping; forensic analysis;

KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;

KM antiinflammatory; gynaecological; reproductive; gene; ss.

OS Homo sapiens.

PN WO200200677-A1.

PD 03-JAN-2002.

PF 07-JUN-2001; 2001MO-US18569.

PR 07-JUN-2000; 2000US-209467P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

PI WPI; 2002-147878/19.

DR P-PSDB; ABP42510.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g.

PT ovarian cancer), immune disorders, cardiovascular disorders and

PT neurological diseases -

PS Claim 1; SEQ ID No 1467; 2922pp; English.

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-

CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen

CC polynucleotides, antibodies against human ovarian antigens, and the use

CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

CC treating, prognosing or preventing various ovary and/or breast-related

CC disorders. Such conditions include ovarian cancer and breast cancer, and

CC metastatic tumours of ovarian or breast origin, reproductive system

disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents cDNA encoding a human ovarian antigen of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 611 BP; 212 A; 106 C; 120 G; 154 T; 19 other;

Query Match 79.0%; Score 15.8; DB 24; Length 611;  
Best Local Similarity 89.5%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0;

OY 2 GTCAAAACCGTGTGTC 20  
||||| ||| |||||  
Db 291 GTCAATACCTTGTGTC 273

Search completed: August 23, 2003, 10:38:17  
Job time : 208.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:18:48 ; Search time 50 Seconds  
(Without alignments)  
176.553 Million cell updates/sec

Title: US-09-927-585a-2

Perfect score: 20

Sequence: 1 tgcataatcggtctactcg 20

Scoring table: IDENTITY\_NTC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*

1: /cgn2\_6/prodata/2/lna/5A\_COMB.seq:\*

2: /cgn2\_6/prodata/2/lna/5B\_COMB.seq:\*

3: /cgn2\_6/prodata/2/lna/6A\_COMB.seq:\*

4: /cgn2\_6/prodata/2/lna/6B\_COMB.seq:\*

5: /cgn2\_6/prodata/2/lna/PCTUS\_COMB.seq:\*

6: /cgn2\_6/prodata/2/lna/backfillseq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20	100.0	30	1	US-08-479-817-9	Sequence 9, Appl1
C 2	20	100.0	30	1	US-08-461-038-9	Sequence 9, Appl1
C 3	20	100.0	30	1	US-08-461-645-9	Sequence 9, Appl1
4	20	100.0	1107	4	US-09-000-094-19	Sequence 19, Appl1
5	20	100.0	1128	4	US-09-000-094-21	Sequence 21, Appl1
6	20	100.0	1398	4	US-09-000-094-23	Sequence 23, Appl1
7	20	100.0	4770	4	US-09-000-094-45	Sequence 45, Appl1
8	20	100.0	8010	4	US-09-521-526-2	Sequence 2, Appl1
9	20	100.0	3340	4	US-09-228-986-5	Sequence 2, Appl1
C 10	15.4	74.0	692	4	US-09-345-2368-76	Sequence 76, Appl1
C 11	14.8	74.0	692	4	US-09-345-2368-78	Sequence 78, Appl1
C 12	14.8	74.0	692	4	US-09-345-2368-80	Sequence 80, Appl1
C 13	14.8	74.0	692	4	US-09-345-2368-82	Sequence 82, Appl1
C 14	14.8	74.0	11970	3	US-09-149-476-80	Sequence 80, Appl1
C 15	14.8	74.0	11970	3	US-09-149-476-82	Sequence 82, Appl1
C 16	14.8	74.0	15231	3	US-09-128-155-16	Sequence 16, Appl1
C 17	14.4	72.0	747	3	US-09-393-877-8	Sequence 8, Appl1
C 18	14.4	72.0	747	3	US-09-393-877-8	Sequence 8, Appl1
C 19	14.2	71.0	747	4	US-08-487-2834-17	Sequence 17, Appl1
C 20	14.2	71.0	747	4	US-08-487-2834-21	Sequence 21, Appl1
C 21	14.2	71.0	747	4	US-08-487-2834-22	Sequence 22, Appl1
C 22	14.2	71.0	747	4	US-08-487-2834-23	Sequence 23, Appl1
C 23	14.2	71.0	750	4	US-08-487-2834-11	Sequence 11, Appl1
C 24	14.2	71.0	750	4	US-08-487-2834-16	Sequence 16, Appl1
C 25	14.2	71.0	820	3	US-08-990-823-54	Sequence 54, Appl1
C 26	14.2	71.0	820	3	US-08-990-823-54	Sequence 54, Appl1
C 27	14.2	71.0	2616	4	US-09-134-001C-1142	Sequence 1142, Ap

28	14.2	71.0	3943	1	US-08-369-796-3	Sequence 3, Appl1
29	14.2	71.0	3943	2	US-08-852-091-3	Sequence 3, Appl1
30	14.2	71.0	3943	2	US-08-820-754-3	Sequence 3, Appl1
31	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
32	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
33	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
34	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
35	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
36	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
37	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appl1
38	14.2	71.0	4003	4	US-09-972-800A-1	Sequence 1, Appl1
39	14.2	71.0	4608	5	US-09-041-886-24	Sequence 24, Appl1
40	14.2	71.0	4608	5	PCT-US94-05277-1	Sequence 1, Appl1
41	14.2	71.0	7739	3	US-09-195-966-1	Sequence 1, Appl1
42	14.2	71.0	7745	1	US-08-299-675-1	Sequence 1, Appl1
43	14.2	71.0	7745	1	US-08-485-241-1	Sequence 1, Appl1
44	14.2	71.0	7745	2	US-08-874-162-1	Sequence 1, Appl1
45	14.2	71.0	34094	4	US-09-292-034-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-08-479-817-9/c  
Sequence 9, Application US/08479817  
Patent No. 5597910  
GENERAL INFORMATION:  
APPLICANT: Gudibande, Satyanarayana R.  
TITLE OF INVENTION: IMPROVED ELECTROCHEMILUMINESCENT LABEL  
TITLE OF INVENTION: FOR DNA PROBE ASSAYS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESSEE: c/o Barry Evans  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,817  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/307,026  
FILING DATE:  
APPLICATION NUMBER: US 07/805,537  
FILING DATE: 11-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Evans, Barry  
REGISTRATION NUMBER: 22,802  
REFERENCE/DOCKET NUMBER: 370068-3440  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-479-817-9  
Query Match 100.0%; Score 20; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATCGGTGCTACCTG 20  
Db 29 TGCTAATCGGTGCTACCTG 10

## RESULT 2

US-08-461-038-9/c  
Sequence 9, Application US/08461038  
Patent No. 5610017  
GENERAL INFORMATION:  
APPLICANT: Gudibande, Satyanarayana R.  
TITLE OF INVENTION: IMPROVED ELECTROCHEMILUMINESCENT LABEL  
TITLE OF INVENTION: FOR DNA PROBE ASSAYS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,038  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Evans, Barry  
REGISTRATION NUMBER: 22,802  
REFERENCE/DOCKET NUMBER: 370068-3451  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-461-038-9

Query Match 100.0%; Score 20; DB 1; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATCGGTGCTACCTG 20  
Db 29 TGCTAATCGGTGCTACCTG 10

## RESULT 3

US-08-461-645-9/c  
Sequence 9, Application US/08461645  
Patent No. 5686244  
GENERAL INFORMATION:  
APPLICANT: Gudibande, Satyanarayana R.  
TITLE OF INVENTION: IMPROVED ELECTROCHEMILUMINESCENT LABEL  
TITLE OF INVENTION: FOR DNA PROBE ASSAYS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,645

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Evans, Barry

REGISTRATION NUMBER: 22,802

REFERENCE/DOCKET NUMBER: 370068-3450

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-461-645-9

Query Match 100.0%; Score 20; DB 1; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.051;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATCGGTGCTACCTG 20  
Db 29 TGCTAATCGGTGCTACCTG 10

## RESULT 4

US-09-000-094-19  
Sequence 19, Application US/09000094  
Patent No. 6365160  
GENERAL INFORMATION:  
APPLICANT: WEBB, Elizabeth Ann  
MARGETTS, Mary Bridgid  
COX, John Cooper  
FRAZER, Ian  
MCWILLIAM, Nigel Alan John  
WILLIAMS, Mark Philip  
MOLONEY, Margaret Bridget  
Holland  
EDWARDS, Stirling John  
TITLE OF INVENTION: PAPILLOMAVIRUS POLYPEPTIDE CONSTRUCTS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/000,094  
FILING DATE: 21-APR-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/AU96/00473  
FILING DATE: 26-JUL-1996  
APPLICATION NUMBER: AU PN 4439/95  
FILING DATE: 27-JUL-1995  
ATTORNEY/AGENT INFORMATION:



NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 017227/0137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1107 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1104  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-000-094-19

Query Match 100.0%; Score 20; DB 4; Length 1107;  
Best Local Similarity 100.0%; Pred. No. 0.082;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGTCGTCTACTG 20  
|||||  
Db 299 TGCTAATTCGTCGTCTACTG 318

RESULT 5  
US-09-000-094-21  
Sequence 21, Application US/09000094  
Patent No. 6365160  
GENERAL INFORMATION:  
APPLICANT: WEBB, Elizabeth Ann  
MARGETTS, Mary Bridgid  
COX, John Cooper  
FRAZER, Ian  
MCWILLAN, Nigel Alan John  
WILLIAMS, Mark Phillip  
MOLONEY, Margaret Bridget  
Holland  
EDWARDS, Stirling John  
TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/000,094  
FILING DATE: 21-Apr-1998  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/AU96/00473  
FILING DATE: 26-JUL-1996  
APPLICATION NUMBER: AU PN 4439/95  
FILING DATE: 27-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 017227/0137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1128 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1125  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-000-094-21

Query Match 100.0%; Score 20; DB 4; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 0.082;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGTCGTCTACTG 20  
|||||  
Db 299 TGCTAATTCGTCGTCTACTG 318

RESULT 6  
US-09-000-094-23  
Sequence 23, Application US/09000094  
Patent No. 6365160  
GENERAL INFORMATION:  
APPLICANT: WEBB, Elizabeth Ann  
MARGETTS, Mary Bridgid  
COX, John Cooper  
FRAZER, Ian  
MCWILLAN, Nigel Alan John  
WILLIAMS, Mark Phillip  
MOLONEY, Margaret Bridget  
Holland  
EDWARDS, Stirling John  
TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/000,094  
FILING DATE: 21-Apr-1998  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/AU96/00473  
FILING DATE: 26-JUL-1996  
APPLICATION NUMBER: AU PN 4439/95  
FILING DATE: 27-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 017227/0137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1398 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:

NAME/KEY: CDS  
LOCATION: 1..1395  
SEQUENCE DESCRIPTION: SEQ ID NO: 23  
US-09-000-094-23

Query Match  
Best Local Similarity 100.0%; Score 20; DB 4; Length 1398;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTGCTACCTG 20  
|||||  
DB 299 TGCTAATTCGGTGCTACCTG 318

RESULT 7  
US-09-000-094-45  
; Sequence 45, Application US/09000094  
; Patent No. 6365160  
; GENERAL INFORMATION:  
; APPLICANT: WEBB, Elizabeth Ann  
; MARGETTS, Mary Bridgid  
; COX, John Cooper  
; FRAZER, Ian  
; MCWILLAN, Nigel Alan John  
; WILLIAMS, Mark Philip  
; MOLONEY, Margaret Bridget  
; Holland  
; EDWARDS, Stirling John  
TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/000,094  
FILING DATE: 21-Apr-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/AU96/00473  
FILING DATE: 26-Jul-1996  
APPLICATION NUMBER: AU PN 4439/95  
FILING DATE: 27-Jul-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 017227/0137  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4770 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4761  
SEQUENCE DESCRIPTION: SEQ ID NO: 45  
US-09-000-094-45

Query Match  
Best Local Similarity 100.0%; Score 20; DB 4; Length 4770;  
Matches 100.0%; Pred. No. 0.099;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTGCTACCTG 20  
|||||  
DB 2357 TGCTAATTCGGTGCTACCTG 2376

RESULT 8  
US-09-521-526-2  
; Sequence 2, Application US/09521526  
; Patent No. 6290965  
; GENERAL INFORMATION:  
; APPLICANT: JANSSEN, KATHRIN U.  
; APPLICANT: HOFMANN, KATHRIN J.  
TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMA VIRUS TYPE  
NUMBER OF SEQUENCES: 6A  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHRISTINE E. CARTY  
STREET: 126 E. LINCOLN AVENUE - P.O. BOX 2000  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/521,526  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/310,468  
FILING DATE: 22-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E.  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19307  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-6734  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8010 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-521-526-2

Query Match  
Best Local Similarity 100.0%; Score 20; DB 3; Length 8010;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTGCTACCTG 20  
|||||  
DB 401 TGCTAATTCGGTGCTACCTG 420

RESULT 9  
PCT-US95-11859-2  
; Sequence 2, Application PC/TUS9511859  
; GENERAL INFORMATION:  
; APPLICANT: JANSSEN, KATHRIN U.  
; APPLICANT: HOFMANN, KATHRIN J.  
TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMA VIRUS TYPE  
NUMBER OF SEQUENCES: 6A  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHRISTINE E. CARTY  
STREET: 126 E. LINCOLN AVENUE - P.O. BOX 2000

CITY: RAHWAY  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/11859  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/310,468  
FILING DATE: 22-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E.  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19307 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-6734  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8010 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
PCT-US95-11859-2

Query Match 100.0%; Score 20; DB 5; Length 8010;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGCTAATCGGTGCTACTG 20  
DB 401 TGCTAATCGGTGCTACTG 420

RESULT 10  
US-09-228-986-5/c  
Sequence 5, Application US/09228986  
Patent No. 6359198  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
FILE REFERENCE: 11000/1020  
CURRENT FILING DATE: 1999-01-12  
CURRENT APPLICATION NUMBER: US/09/228,986  
NUMBER OF SEQ ID NOS: 130  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 3340  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-228-986-5

Query Match 77.0%; Score 15.4; DB 4; Length 3340;  
Best Local Similarity 94.1%; Pred. No. 31;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GCTAATCGGTGCTACC 18  
DB 1176 GCTAATCGGTGCTACC 1160

RESULT 11  
US-09-345-236B-76  
Sequence 76, Application US/09345236B

Patent No. 6521454  
GENERAL INFORMATION:  
APPLICANT: Becnel, James J.  
APPLICANT: Tukuo, Fukuda  
APPLICANT: Moser, Bettina  
APPLICANT: Cockburn, Andrew  
APPLICANT: White, Susan E.  
APPLICANT: Undeen, Albert H.  
TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal  
FILE REFERENCE: 21042.0004  
CURRENT FILING DATE: 1999-06-30  
CURRENT APPLICATION NUMBER: US/09/345,236B  
NUMBER OF SEQ ID NOS: 148  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 76  
LENGTH: 692  
TYPE: DNA  
ORGANISM: mosquito baculovirus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (27)...(689)  
US-09-345-236B-76

Query Match 74.0%; Score 14.8; DB 4; Length 692;  
Best Local Similarity 88.9%; Pred. No. 54;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTAATTCGGTCTACTG 20  
DB 165 CTAATTCGGTCTACTG 182

RESULT 12  
US-09-345-236B-78  
Sequence 78, Application US/09345236B  
Patent No. 6521454  
GENERAL INFORMATION:  
APPLICANT: Becnel, James J.  
APPLICANT: Tukuo, Fukuda  
APPLICANT: Moser, Bettina  
APPLICANT: Cockburn, Andrew  
APPLICANT: White, Susan E.  
APPLICANT: Undeen, Albert H.  
TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal  
FILE REFERENCE: 21042.0004  
CURRENT FILING DATE: 1999-06-30  
CURRENT APPLICATION NUMBER: US/09/345,236B  
NUMBER OF SEQ ID NOS: 148  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 78  
LENGTH: 692  
TYPE: DNA  
ORGANISM: mosquito baculovirus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (13)...(165)  
US-09-345-236B-78

Query Match 74.0%; Score 14.8; DB 4; Length 692;  
Best Local Similarity 88.9%; Pred. No. 54;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTAATTCGGTCTACTG 20  
DB 165 CTAATTCGGTCTACTG 182

RESULT 13  
US-09-345-236B-80/c  
Sequence 80, Application US/09345236B  
Patent No. 6521454

GENERAL INFORMATION:  
APPLICANT: Becnel, James J.  
APPLICANT: Tokuo, Fukuda  
APPLICANT: Moser, Bettina  
APPLICANT: Cockburn, Andrew  
APPLICANT: White, Susan E.  
APPLICANT: Undeen, Albert H.  
TITLE OF INVENTION: No. 6521454e1 Baculoviruses, Insecticidal  
TITLE OF INVENTION: Compositions, and Methods for Control of Invertebrates  
FILE REFERENCE: 21042.0004  
CURRENT APPLICATION NUMBER: US/09/345,236B  
NUMBER OF SEQ ID NOS: 148  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 80  
LENGTH: 692  
TYPE: DNA  
ORGANISM: mosquito baculovirus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(366)  
NAME/KEY: CDS  
LOCATION: (529)...(690)  
US-09-345-236B-80

Query Match 74.0%; Score 14.8; DB 4; Length 692;  
Best Local Similarity 88.9%; Pred. No. 54;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CTAATCGCTGCTACTG 20  
DB 528 CTAATCGCTGCTGCTG 511

RESULT 14  
US-09-149-476-80  
Sequence 80, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,882  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,637  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,903  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,888  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,879  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,880  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,894  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,911  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,636  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,874  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,910  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,845  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,892  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/057,761  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/047,595  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,599  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,588  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,585  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,586  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,590  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,594  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,589  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,593  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/047,614  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/043,576  
 EARLIER FILING DATE: 1997-04-11  
 EARLIER APPLICATION NUMBER: 60/043,576  
 EARLIER FILING DATE: 1997-04-11  
 EARLIER APPLICATION NUMBER: 60/047,501  
 EARLIER FILING DATE: 1997-05-23  
 EARLIER APPLICATION NUMBER: 60/043,670  
 EARLIER FILING DATE: 1997-04-11  
 EARLIER APPLICATION NUMBER: 60/056,632  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,664  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,876  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,881  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,909  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,875  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,862

EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,887  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/056,908  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/048,964  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/057,650  
 EARLIER FILING DATE: 1997-09-05  
 EARLIER APPLICATION NUMBER: 60/056,884  
 EARLIER FILING DATE: 1997-08-22  
 EARLIER APPLICATION NUMBER: 60/057,669  
 EARLIER FILING DATE: 1997-09-05  
 EARLIER APPLICATION NUMBER: 60/049,610  
 EARLIER FILING DATE: 1997-06-13  
 EARLIER APPLICATION NUMBER: 60/061,060  
 EARLIER FILING DATE: 1997-10-02

Query Match 74.0%; Score 14.8; DB 4; Length 1285;  
 Best Local Similarity 88.9%; Pred. No. 58;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTAATCGGTCTACCTG 20  
 Db 419 CTAATCTGTCTCCCTG 436

RESULT 15  
 US-09-345-217-1  
 : Sequence 1, Application US/09345217  
 : Patent No. 6268142  
 : GENERAL INFORMATION:  
 : APPLICANT: DUFE, GORDON W.  
 : APPLICANT: COX, ANGELA  
 : APPLICANT: CAMP, NICOLA J.  
 : TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR DISEASES ASSOCIATED  
 : TITLE OF INVENTION: WITH AN IL-1 INFLAMMATORY HAPLOTYPE  
 : FILE REFERENCE: MSA-010.02  
 : CURRENT APPLICATION NUMBER: US/09/345,217  
 : EARLIER FILING DATE: 1999-06-30  
 : EARLIER APPLICATION NUMBER: PCT/GB98/01481  
 : EARLIER FILING DATE: 1998-05-21  
 : EARLIER APPLICATION NUMBER: 9711040.7  
 : EARLIER FILING DATE: 1997-05-29  
 : NUMBER OF SEQ ID NOS: 32  
 : SOFTWARE: Patent Ver. 2.0  
 : SEQ ID NO 1  
 : LENGTH: 11970  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 US-09-345-217-1

Query Match 74.0%; Score 14.8; DB 3; Length 11970;  
 Best Local Similarity 88.9%; Pred. No. 78;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 3 CTAATCGGTCTACCTG 20  
 Db 5364 CTAATCTGTCTCCCTG 5381

Search completed: August 23, 2003, 12:15:10  
 Job time : 52 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:14:58 ; Search time 1848 Seconds  
(Without alignments)  
263.035 Million cell updates/sec

Title: US-09-927-585a-2

Perfect score: 20  
Sequence: 1 tgcattcgcgtctactctg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estln:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estl:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pmg:\*  
27: gb\_gss1:\*  
28: gb\_gss2:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	17.4	87.0	29	CC149256 ZMMBB15H
2	16.8	84.0	10	BE436705 BST407823
3	16.8	84.0	28	AQ137540 HS_2252_A
4	16.8	84.0	29	CNS04H07 Tetraodon

Result No.	Score	Query Length	ID	Description
5	16.4	82.0	29	CC122092
6	16.4	82.0	29	CC134974
7	16.4	82.0	29	CC136978
8	15.8	79.0	226	BB570973
9	15.8	79.0	273	BB878622
10	15.8	79.0	300	AU114943
11	15.8	79.0	376	BB397091
12	15.8	79.0	400	CC159313
13	15.8	79.0	427	BF937261
14	15.8	79.0	470	AQ678663
15	15.8	79.0	523	BQ378925
16	15.8	79.0	532	BQ058697
17	15.8	79.0	536	CD437384
18	15.8	79.0	540	AQ766138
19	15.8	79.0	543	BQ077290
20	15.8	79.0	545	CB081561
21	15.8	79.0	551	B1984033
22	15.8	79.0	575	AQ491873
23	15.8	79.0	577	BM192517
24	15.8	79.0	586	CB879626
25	15.8	79.0	588	BQ071409
26	15.8	79.0	635	BZ067787
27	15.8	79.0	646	AZ082092
28	15.8	79.0	670	CC089529
29	15.8	79.0	710	CC129477
30	15.8	79.0	740	CC129477
31	15.8	79.0	743	BM407709
32	15.8	79.0	782	BZ698014
33	15.8	79.0	786	BZ698353
34	15.8	79.0	825	BZ677257
35	15.8	79.0	826	BE884306
36	15.8	79.0	855	CC066386
37	15.8	79.0	869	BQ734750
38	15.8	79.0	890	BQ730599
39	15.8	79.0	925	CC079842
40	15.8	79.0	948	CNS020NR
41	15.8	79.0	1052	CC224089
42	15.8	79.0	1228	CC232880
43	15.4	77.0	237	AI335576
44	15.4	77.0	238	BQ345302
45	15.4	77.0	304	AA188347

## ALIGNMENTS

RESULT 1  
CC149256 887 bp DNA linear GSS 24-APR-2003  
ZMMBB15H20.r ZMMBB15H20  
DEFINITION  
3', genomic survey sequence.

ACCESSION  
CC149256  
VERSION  
CC149256.1 GI:30094449

KEYWORDS  
GSS.

SOURCE  
Zea mays subsp. mays (maize)

ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
1 (bases 1 to 887)  
Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bhatti, A.K., Messing, J.

AUTHORS  
and Wang, R.

TITLE  
Sequencing of the maize genome

JOURNAL  
Unpublished

COMMENT  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA

Rel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers





SOURCE Tetraodon nigroviridis  
ORGANISM Tetraodon nigroviridis

REFERENCE 1  
AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Mincker, P., Brottier, P., Quetier, F., Saurin, W., and Weissenbach, J.  
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)  
MEDLINE 20296633  
PUBMED 10835645

REFERENCE 2  
AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.  
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
JOURNAL Genome Res. 10 (7), 939-949 (2000)  
MEDLINE 20359837  
PUBMED 10899143

REFERENCE 3 (bases 1 to 1096)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES  
source Location/Qualifiers  
1..1096  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone\_1lb="109D19"  
/clone\_1lb="g"  
/note="Genoscope sequence ID : COBG109CH105P1-end : POC-Or1"

BASE COUNT 310 a 232 c 232 g 316 t 6 others  
ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 1096;  
Best Local Similarity 90.0%; Pred. No. 8e+02; Mismatches 2; Indels 0; Gaps 0;  
Matches 18; Conservative 0;

QY 1 TCCTAATTCGTCCTACTG 20  
||||| ||||||| |||  
Db 724 TCCTAATTCGTCCTACTG 743

RESULT 5  
LOCUS CC122092 600 bp DNA linear GSS 16-APR-2003  
DEFINITION ND1.46E23.SP6 Notre Dame Liverpool Aedes aegypti genomic clone  
ACCESSION CC122092  
VERSION CC122092.1 GI:29991147  
KEYWORDS GSS.  
SOURCE Aedes aegypti (yellow fever mosquito)  
ORGANISM Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.  
1 (bases 1 to 600)  
Loftus, B., Shetty, J., Knudson, D., and Severson, D.  
BAC end sequencing of Aedes aegypti  
Unpublished  
Other\_GSSs: ND1.46E23.T7  
COMMENT Contact: Brendan Loftus

Department of Eukaryotic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: entaetigr.org  
Library was provided by David Severson  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..600  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="Liverpool"  
/db\_xref="taxon:7159"  
/clone\_1lb="ND1.46E23"  
/clone\_1lb="Notre Dame Liverpool"  
/note="Vector: pECBAC1, Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

BASE COUNT 180 a 160 c 181 g 79 t  
ORIGIN

Query Match 82.0%; Score 16.4; DB 29; Length 600;  
Best Local Similarity 94.4%; Pred. No. 1.1e+03; Mismatches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Matches 17; Conservative 0;

QY 1 TCCTAATTCGTCCTACC 18  
||||| ||||||| |||  
Db 349 TCCTAATTCGTCCTACC 332

RESULT 6  
LOCUS CC134974 854 bp DNA linear GSS 16-APR-2003  
DEFINITION ND1.91K13.SP6 Notre Dame Liverpool Aedes aegypti genomic clone  
ACCESSION CC134974  
VERSION CC134974.1 GI:30004029  
KEYWORDS GSS.  
SOURCE Aedes aegypti (yellow fever mosquito)  
ORGANISM Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.  
1 (bases 1 to 854)  
Loftus, B., Shetty, J., Knudson, D., and Severson, D.  
BAC end sequencing of Aedes aegypti  
Unpublished  
Other\_GSSs: ND1.91K13.T7  
Contact: Brendan Loftus  
Department of Eukaryotic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: entaetigr.org  
Library was provided by David Severson  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..854  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="Liverpool"  
/db\_xref="taxon:7159"  
/clone\_1lb="ND1.91K13"  
/clone\_1lb="Notre Dame Liverpool"  
/note="Vector: pECBAC1, Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

BASE COUNT 239 a 208 c 240 g 167 t

Query Match	82.0%;	Score 16.4;	DB 29;	Length 854;
Best Local Similarity	94.4%;	Pred. No. 1.2e+03;		
Matches 17; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	1	TGCTAATTGCGTGTCTACC	18	
Db	399	TGCTAATTGCGTGTCTACC	382	
RESULT 7				
CC136978/c		932 bp	DNA	linear
DEFINITION	NDL.91B22.sp6 Notre Dame Liverpool Aedes aegypti genomic clone			
ACCESSION	NDL.91B22. genomic survey sequence.			
VERSION	CC136978			
KEYWORDS	CC136978.1 GI:30006033			
SOURCE	GSS.			
ORGANISM	Aedes aegypti (yellow fever mosquito)			
REFERENCE	Aedes aegypti			
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes. 1 (bases 1 to 932)			
TITLE	Loftus, B., Shetty, J., Knudson, D. and Severson, D.			
JOURNAL	BAC end sequencing of Aedes aegypti unpublished			
COMMENT	Other_GSSs: NDL.91B22.T7			
	Contact: Brendan Loftus			
	Department of Eukaryotic Genomics			
	TIGR			
	9712 Medical Center Drive, Rockville, MD 20850, USA			
	Tel: 301-838-3543			
	Fax: 301-838-0208			
	Email: entae@tigr.org			
	Library was provided by David Severson			
	Seq primer: SP6			
	Class: BAC ends.			
FEATURES				
source	Location/Qualifiers			
	1..932			
	/organism="Aedes aegypti"			
	/mol_type="genomic DNA"			
	/strain="Liverpool"			
	/db_xref="taxon:7159"			
	/clone="NDL.91B22"			
	/clone_11b="Notre Dame Liverpool"			
	/note="Vector: pGCBAC1; Site_1: Hind III; The library was prepared from whole body tissue of newly hatched 1d larvae by David Severson at the University of Notre Dame and Hongbin Zhang"			
BASE COUNT	262 a 221 c 265 g 184 t			
ORIGIN				
Query Match	82.0%;	Score 16.4;	DB 29;	Length 932;
Best Local Similarity	94.4%;	Pred. No. 1.2e+03;		
Matches 17; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	1	TGCTAATTGCGTGTCTACC	18	
Db	396	TGCTAATTGCGTGTCTACC	379	
RESULT 8				
BB570973		226 bp	mRNA	linear
DEFINITION	BB570973 RIKEN full-length enriched, 0 day neonate skin Mus			
ACCESSION	musculus cDNA clone 4633401D12 5', mRNA sequence.			
VERSION	BB570973			
KEYWORDS	BB570973.1 GI:11461881			
SOURCE	EST.			
ORGANISM	Mus musculus (house mouse)			
	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 226)
TITLE JOURNAL COMMENT	Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanganaki, T., Hayatsu, N., Hirooka, T., Hirozane, T., Hodayama, Y., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno, N., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shingara, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toyota, T., Wataniki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Aizawa, K. et al. 2000) Unpublished Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoke, S., Sasakita, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2). 520-524 (1998) Itoh, M., Kitsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5). 463-470 (1999) Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site ( <a href="http://genome.irc.riken.go.jp">http://genome.irc.riken.go.jp</a> ) for further details.
FEATURES SOURCE	location/Qualifiers 1..226 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /cclone="463340D12" /sex="mixed" /tissue_type="skin" /dev_stage="0 day neonate" /lab_host="DH10B" /cclone_lib="RIKEN full-length enriched, 0 day neonate skin" note=Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group In Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAAGAGATCCAAAGACCTCTTTTTCCTTTTTTNN 3'], cDNA was prepared by using trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAAGATCCGAAGTATTATAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pluscript KS(+) after bulk excision from Lambda Phix I"
BASE COUNT ORIGIN	63 a                 56 c                 52 g                 55 t  Query Match                 79.0%; Score 15.8; DB 10; Length 226; Best Local Similarity       89.5%; Pred. No. 1.66+03; Matches       17; Conservative       0; Mismatches       2; Indels       0; Gaps       0;

QY 1 TGCTAATTCGCTGCTACT 19  
 DB 174 TTCTAATTAGTGTCTACT 192  
 RESULT 9  
 LOCUS BH878622/c  
 DEFINITION bh878622 273 bp DNA linear GSS 05-AUG-2002  
 genomic clone hs83c03 5', genomic survey sequence.  
 ACCSSION BH878622  
 VERSION BH878622.1 GI:22114519  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 273)  
 Rabinowicz,P.D., O'Shaughnessy,A.L., Ballja,V., Dedhia,N.,  
 Katzenburger,F., King,L., Miller,B., Miller,S., Nascimento,L.,  
 Zlatavern,T., McCombie,W.R. and Martienssen,R.A.,  
 Genomic Shotgun sequences from Zea mays (methyl-filtered)  
 Unpublished  
 CONTACT: W. Richard McCombie  
 Lita Annenberg Hazen Genome Sequencing Center  
 Cold Spring Harbor Laboratory  
 PO Box 100, Cold Spring Harbor, NY 11724, USA  
 Tel: 516 367 8884  
 Fax: 516 367 8874  
 Email: mcombie@cshl.org  
 Plate: hs83 row: c column: 03  
 Seq primer: -21M13univfw  
 Class: shotgun  
 High quality sequence stop: 273.  
 FEATURES  
 Location/Qualifiers  
 1..273  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultiyar="B73"  
 /db\_xref="taxon:4577"  
 /clone="hs83c03"  
 /lab\_host="JM107 or DH5a"  
 /note="Organ: Immature ears; Site\_1: Xba I; Site\_2: Xba I;  
 The vector was digested with XbaI and one nucleotide was  
 added by fill in in the recessive 3' end. The genomic DNA  
 was nebulized, end repaired, adaptor ligated and size  
 fractionated using sephadex. The resulting fragments were  
 between 0.8 and 3 kb and were cloned into the vector  
 (x/y reads in M13mp19, b/g reads in pUC19). The same  
 ligation was transformed in either JM107 or DH5a."  
 BASE COUNT 78 a 58 c 47 g 90 t  
 ORIGIN  
 Query Match 79.0%; Score 15.8; DB 28; Length 273;  
 Best Local Similarity 89.5%; Pred. No. 1.7e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GCTAATTCGGTGTACTCG 20  
 DB 44 GCAATTCAGTGTCTACTG 26  
 RESULT 10  
 LOCUS AU114943/c  
 DEFINITION AU114943 300 bp mRNA linear EST 19-OCT-2000  
 elegans cDNA clone yk723h1 3', mRNA sequence.  
 ACCSSION AU114943  
 VERSION AU114943.1 GI:10928510  
 KEYWORDS EST.

SOURCE Caenorhabditis elegans  
 ORGANISM Caenorhabditis elegans  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
 ; Rhabditidae; Peloderinae; Caenorhabditis.  
 REFERENCE 1 (bases 1 to 300)  
 Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
 and Sugano,S.  
 A complementary view of the C. elegans genome  
 Unpublished  
 CONTACT: Yuji Kohara  
 Genome Biology Lab.  
 National Institute of Genetics  
 Yata 1111, Mishima, Shizuoka 411, Japan  
 Tel: 81-559-81-6854  
 Fax: 81-559-81-6855  
 Email: ykohara@lab.nig.ac.jp.  
 FEATURES  
 Location/Qualifiers  
 1..300  
 /organism="Caenorhabditis elegans"  
 /mol\_type="mRNA"  
 /strain="N2"  
 /db\_xref="taxon:6239"  
 /clone="yk723h1"  
 /sex="Hermaphrodite"  
 /tissue\_type="whole animal"  
 /dev\_stage="varied"  
 /clone\_lib="unpublished oligo-capped cDNA library"  
 BASE COUNT 95 a 62 c 46 g 97 t  
 ORIGIN  
 Query Match 79.0%; Score 15.8; DB 9; Length 300;  
 Best Local Similarity 89.5%; Pred. No. 1.7e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TGCTAATTCGCTGCTACT 19  
 DB 261 TACTAATTCGCTGCTACT 243  
 RESULT 11  
 LOCUS BH397091/c  
 DEFINITION AG-ND-168H24.TF ND-TAM Anopheles gambiae genomic clone AG-ND-168H24  
 , genomic survey sequence.  
 ACCSSION BH397091  
 VERSION BH397091.1 GI:17343307  
 KEYWORDS GSS.  
 SOURCE Anopheles gambiae (African malaria mosquito)  
 ORGANISM Anopheles gambiae  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae.  
 REFERENCE 1 (bases 1 to 376)  
 Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., Ren  
 ,C., Huff,E.R., Carille,J.L., Black,K., Zhang,H.-B., Gardner,M.J.  
 and Collins,F.H.  
 Construction of a BAC library and generation of BAC end  
 sequence-tagged connectors for genome sequencing of the African  
 malaria mosquito Anopheles gambiae  
 Mol. Genet. Genomics 268 (6), 720-728 (2003)  
 MEDLINE 22542063  
 PUBMED 12655398  
 COMMENT Other\_GSSs: AG-ND-168H24.TR  
 Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: b.loftus@tigr.org  
 This clone is from an A. gambiae BAC library (ND-TAM) provided by  
 F.H. Collins and sequenced by The Institute for Genomic Research  
 (TIGR). The BAC library was generated from A. gambiae PEST strain

DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.

Seq primer: M3 For  
Class: BAC ends.

FEATURES  
source Location/Qualifiers

1..376  
/organism="Anopheles gambiae"  
/mol\_type="genomic DNA"  
/strain="PEST"  
/db\_xref="taxon:7165"  
/clone="AG-ND-168H24"  
/clone\_1lb="ND-TAM"  
/note="Vector: pECBAC1; Site\_1: HindIII"

BASE COUNT 108 a 75 c 101 g 92 t  
ORIGIN

Query Match 79.0%; Score 15.8; DB 28; Length 376;  
Best Local Similarity 89.5%; Pred. No. 1.8e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCGTATTCGGCTACCT 19  
Db 248 TCGTATTCGGCTACCT 230

RESULT 12  
CC159313/C 400 bp DNA linear GSS 29-APR-2003  
LOCUS 1901C01.b1 WGS-Zmaysf (DH5a methyl filtered) Zea mays genomic clone  
DEFINITION CC159313  
ACCESSION CC159313  
VERSION CC159313.1 GI:30184091  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 400)

REFERENCE  
AUTHORS Rabinowitz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L., Zutavern, T., McCombie, W.R. and Martienssen, R.A.  
Genomic shotgun sequences from Zea mays (methyl-filtered)  
Unpublished

CONTACT: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cs.cshl.org

Plate: 1901 row: C column: 01  
Seq primer: -21M3univFwd  
Class: Shotgun  
High quality sequence stop: 400.

FEATURES  
source Location/Qualifiers

1..400  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="1901C01"  
/lab\_host="DH5a"

/clone\_1lb="WGS-Zmaysf (DH5a methyl filtered)"  
/note="Organ: immature ears; Site\_1: Xba I; Site\_2: Xba I;  
The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were

between 0.8 and 3 kb and were cloned into the vector (x/y reads in M3imp19...b/g reads in pUC19). The same ligation was transformed into DH5a."

BASE COUNT 128 a 73 c 72 g 127 t

Query Match 79.0%; Score 15.8; DB 29; Length 400;  
Best Local Similarity 89.5%; Pred. No. 1.9e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTATTCGGCTACCTG 20  
Db 210 GCAATTCAGTCTACTG 192

RESULT 13  
BF937261 427 bp mRNA linear EST 13-FEB-2002  
LOCUS fm62g01.y1 zebrafish adult retina cdna Danio rerio cdna clone  
DEFINITION IMAGE:4200001 5', mRNA sequence.  
ACCESSION BF937261  
VERSION BF937261.1 GI:12354581  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)

Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 427)  
REFERENCE  
AUTHORS Clark, M., Johnson, S.L., Lehnach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theisling, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

Washu Zebrafish EST Project 1998

Unpublished

CONTACT: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: zbrfish@wustl.edu  
Library constructed by: Chandra Tucker and Gregory Niemi DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Ressourcenzentrum/Primatdatenbank, Berlin, Germany (web address: www.rzpd.de)  
Seq primer: T3 ET from Amersham

High quality sequence stop: 421.

FEATURES  
source Location/Qualifiers

1..427  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/strain="wild-type"  
/db\_xref="taxon:7055"  
/clone="IMAGE:4200001"  
/sex="mixed"  
/dev\_stage="1-2 years"  
/lab\_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"

/clone\_1lb="zebrafish adult retina cdna"  
/note="Vector: Lambda ZAP II (pluscript SK-); Site\_1: EcoRI; Site\_2: SalI; This zebrafish library was constructed by Dr. Susan E. Brockerhoff (email: sbrocker@u.washington.edu) RZPD library number: 760"

BASE COUNT 135 a 106 c 97 g 89 t

Query Match 79.0%; Score 15.8; DB 10; Length 427;  
Best Local Similarity 89.5%; Pred. No. 1.9e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTATTCGGCTACTG 20  
||||||| ||||| |||

Db 310 GCTAATTCAGTCTACTG 328

RESULT 14  
LOCUS A0678663

DEFINITION A0678663 470 bp DNA linear GSS 25-JUN-1999  
HS\_2094\_B2\_F10\_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2094 COL-20 Row-L, genomic survey sequence.

ACCESSION A0678663  
VERSION A0678663.1 GI:5227467  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 470)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
BAC ends may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: <http://www.htsc.washington.edu>  
Plate: 2094 Row: L Column: 20  
Seq primer: M13 Reverse  
Class: BAC ends  
High quality sequence stop: 470.  
Location/Qualifiers  
1..470  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=2094 Col=20 Row=L"  
/sex="male"  
/note="Organ: Sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 123 a 100 c 79 g 168 t

ORIGIN

Query Match 79.0%; Score 15.8; DB 28; Length 470;  
Best Local Similarity 89.5%; Pred. No. 2e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCTAATTCGCTGCTACT 19  
|||||  
Db 252 TGCTGATTCAGTCTACT 270

RESULT 15  
LOCUS BQ378925 523 bp mRNA linear EST 21-MAY-2002  
DEFINITION RC2-UT0087-200900-110-h04 UT0087 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BQ378925  
VERSION BQ378925.1 GI:21054439  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 523)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2704922  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC2&t=RC2-UT0087-200900-110-h04&t3=2000-09-20&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 21  
High quality sequence stop: 523.  
Location/Qualifiers  
1..523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone.lib="UT0087"  
/note="Organ: uterus; tumor: Vector: puc18; Site: 1; Smat: Site: 2; Smat: A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 163 a 109 c 99 g 152 t

ORIGIN

Query Match 79.0%; Score 15.8; DB 13; Length 523;  
Best Local Similarity 89.5%; Pred. No. 2e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTAATTCGCTGCTACTG 20  
|||||  
Db 448 GCAAAATTCGCTGCTACTG 430

Search completed: August 23, 2003, 12:13:24  
Job time: 1852 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 11:11:44 ; Search time 141 Seconds  
(without alignments)  
318.888 Million cell updates/sec

Title: US-09-927-585a-2

Perfect score: 20

Sequence: 1 tgcattcgcgtgctactg 20

Scoring table: IDENTITY\_NUC

Searched: 1517243 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCr\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCrS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	85.0	427	13	US-10-027-632-74539
2	15.8	79.0	281	9	US-09-764-878-249
3	15.8	79.0	281	9	US-09-764-860-249
4	15.8	79.0	281	14	US-10-079-854-249
5	15.8	79.0	281	14	US-10-074-095-928
6	15.8	79.0	287	9	US-09-764-878-36
7	15.8	79.0	287	9	US-09-764-860-170
8	15.8	79.0	287	14	US-10-079-854-36
9	15.8	79.0	287	14	US-10-074-095-170
10	15.8	79.0	4682	12	US-10-128-714-257
11	15.8	79.0	14468	14	US-10-295-362-19
12	15.4	77.0	3340	14	US-10-101-464A-5
13	15.2	76.0	245	13	US-10-044-080-752
14	15.2	76.0	457	13	US-10-027-632-35375
15	15.2	76.0	457	13	US-10-027-632-60413
16	15.2	76.0	457	13	US-10-027-632-298306

C 17	15.2	76.0	611	13	US-10-027-632-198643	Sequence 198643,
C 18	15.2	76.0	611	13	US-10-027-632-198644	Sequence 198644,
C 19	15.2	76.0	795	13	US-10-027-632-165367	Sequence 165367,
C 20	15	75.0	675	13	US-10-027-632-262303	Sequence 262303,
C 21	14.8	74.0	451	11	US-09-918-995-12873	Sequence 12873, A
C 22	14.8	74.0	654	13	US-10-027-632-243504	Sequence 243504,
C 23	14.8	74.0	690	13	US-10-027-632-27924	Sequence 27924, A
C 24	14.8	74.0	690	13	US-10-027-632-27925	Sequence 27925, A
C 25	14.8	74.0	690	13	US-10-027-632-27926	Sequence 27926, A
C 26	14.8	74.0	1158	13	US-10-027-632-216899	Sequence 216899,
C 27	14.8	74.0	1285	11	US-09-809-391-80	Sequence 80, Appl
C 28	14.8	74.0	2029	13	US-10-027-632-264047	Sequence 264047,
C 29	14.8	74.0	3334	13	US-10-108-605-316	Sequence 316, App
C 30	14.8	74.0	7036	14	US-10-178-782-3	Sequence 3, Appl1
C 31	14.8	74.0	8326	11	US-09-764-872-716	Sequence 716, App
C 32	14.8	74.0	11970	10	US-09-845-129-1	Sequence 1, Appl1
C 33	14.8	74.0	11970	12	US-10-172-919-1	Sequence 1, Appl1
C 34	14.8	74.0	11970	14	US-10-167-127-1	Sequence 717, App
C 35	14.8	74.0	21732	11	US-09-764-872-717	Sequence 16, Appl
C 36	14.8	74.0	152331	13	US-10-095-407-16	Sequence 83, Appl
C 37	14.4	72.0	36	11	US-09-845-616-83	Sequence 203394,
C 38	14.4	72.0	650	13	US-10-027-632-202394	Sequence 202394,
C 39	14.4	72.0	650	13	US-10-027-632-202395	Sequence 160032,
C 40	14.4	72.0	674	13	US-10-027-632-160032	Sequence 160033,
C 41	14.4	72.0	674	13	US-10-027-632-160033	Sequence 252549,
C 42	14.4	72.0	1348	13	US-10-027-632-252549	Sequence 252550,
C 43	14.4	72.0	1348	13	US-10-027-632-252550	Sequence 60, Appl
C 44	14.4	72.0	4575	10	US-09-965-703-60	Sequence 49, Appl
C 45	14.2	71.0	355	9	US-09-813-358-49	

## ALIGNMENTS

RESULT 1  
US-10-027-632-74539  
Sequence 74539, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: Polymorphisms in the Human Genome  
CURRENT APPLICATION NUMBER: 108827.129  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 74539  
LENGTH: 427  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-74539

Query Match 85.0%; Score 17; DB 13; Length 427;  
Best Local Similarity 89.5%; Pred. No. 2;  
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GCTAATTCGCTGCTACTG 20  
Db 276 GYTAATTCGCTGCTACTG 294

```

RESULT 2
US-09-764-878-249
; Sequence 249, Application US/09764878
; Patent No. US20020090515A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121
; CURRENT APPLICATION NUMBER: US/09/764,878
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 249
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-249

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 281;
Best Local Similarity 89.5%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TGCTAATTCGCTGCTACT 19
    ||||| ||| ||||| |||
Db 102 TGCTAATTCGCTGCTACT 120

```

```

RESULT 3
US-09-764-860-928
; Sequence 928, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 928
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-928

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 281;
Best Local Similarity 89.5%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TGCTAATTCGCTGCTACT 19
    ||||| ||| ||||| |||
Db 102 TGCTAATTCGCTGCTACT 120

```

```

RESULT 4
US-10-079-854-249
; Sequence 249, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; PRIORITY FILING DATE: 2002-02-22
; Prior application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 249
; LENGTH: 281

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-079-854-249

```

```

Query Match          79.0%; Score 15.8; DB 14; Length 281;
Best Local Similarity 89.5%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TGCTAATTCGCTGCTACT 19
    ||||| ||| ||||| |||
Db 102 TGCTAATTCGCTGCTACT 120

```

```

RESULT 5
US-10-074-095-928
; Sequence 928, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; PRIORITY FILING DATE: 2002-02-14
; Prior application number: 09/764,860
; PRIORITY FILING DATE: 2001-01-17
; Prior application number: 60/179,065
; PRIORITY FILING DATE: 2000-01-31
; Prior application number: 60/180,628
; PRIORITY FILING DATE: 2000-02-04
; Prior application number: 60/214,886
; PRIORITY FILING DATE: 2000-06-28
; Prior application number: 60/217,487
; PRIORITY FILING DATE: 2000-07-11
; Prior application number: 60/225,758
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/220,963
; PRIORITY FILING DATE: 2000-07-26
; Prior application number: 60/217,496
; PRIORITY FILING DATE: 2000-07-11
; Prior application number: 60/225,447
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/218,290
; PRIORITY FILING DATE: 2000-07-14
; Prior application number: 60/225,757
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/226,868
; PRIORITY FILING DATE: 2000-08-22
; Prior application number: 60/216,647
; PRIORITY FILING DATE: 2000-07-07
; Prior application number: 60/225,267
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/216,880
; PRIORITY FILING DATE: 2000-07-07
; Prior application number: 60/225,270
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/225,869
; PRIORITY FILING DATE: 2000-12-08
; Prior application number: 60/235,834
; PRIORITY FILING DATE: 2000-09-27
; Prior application number: 60/234,274
; PRIORITY FILING DATE: 2000-09-21
; Prior application number: 60/234,223
; PRIORITY FILING DATE: 2000-09-21
; Prior application number: 60/228,924
; PRIORITY FILING DATE: 2000-08-30
; Prior application number: 60/224,518
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/236,369
; PRIORITY FILING DATE: 2000-09-29
; Prior application number: 60/224,519
; PRIORITY FILING DATE: 2000-08-14
; Prior application number: 60/220,964
; PRIORITY FILING DATE: 2000-07-26

```



PRIOR APPLICATION NUMBER: 60/241,809  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/249,299  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/236,327  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/241,785  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/244,617  
 PRIOR FILING DATE: 2000-11-01  
 PRIOR APPLICATION NUMBER: 60/225,268  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/236,368  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/251,856  
 PRIOR FILING DATE: 2000-12-08  
 PRIOR APPLICATION NUMBER: 60/251,868  
 PRIOR FILING DATE: 2000-12-08  
 PRIOR APPLICATION NUMBER: 60/229,344  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/234,997  
 PRIOR FILING DATE: 2000-09-25  
 PRIOR APPLICATION NUMBER: 60/229,343  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/229,345  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/229,287  
 PRIOR FILING DATE: 2000-09-01  
 PRIOR APPLICATION NUMBER: 60/229,513  
 PRIOR FILING DATE: 2000-09-05  
 PRIOR APPLICATION NUMBER: 60/231,413  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/229,509  
 PRIOR FILING DATE: 2000-09-05  
 PRIOR APPLICATION NUMBER: 60/236,367  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/237,039  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/237,038  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/236,370  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: 60/236,802  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/237,037  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/237,040  
 PRIOR FILING DATE: 2000-10-02  
 PRIOR APPLICATION NUMBER: 60/240,960  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/239,935  
 PRIOR FILING DATE: 2000-10-13  
 PRIOR APPLICATION NUMBER: 60/239,937  
 PRIOR FILING DATE: 2000-10-13  
 PRIOR APPLICATION NUMBER: 60/241,787  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,474  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/246,532  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/249,216  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,210  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/226,681  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,759  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/225,213  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/227,182  
 PRIOR FILING DATE: 2000-08-22  
 PRIOR APPLICATION NUMBER: 60/225,214

PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/235,836  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: 60/230,438  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/215,135  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: 60/225,266  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/249,218  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,208  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,213  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,212  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,207  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,245  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,244  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,217  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,211  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,215  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,264  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/231,242  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,081  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,080  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,414  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,244  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/233,064  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/233,063  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,397  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,399  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,401  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/241,808  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,826  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,786  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,221  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,475  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/231,243  
 PRIOR FILING DATE: 2000-09-08

Query Match 79.0%; Score 15.8; DB 14; Length 281;  
 Best Local Similarity 89.5%; Pred. No. 1e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTACT 19  
|||||  
Db 102 TGCTAATTCGTGTACT 120

## RESULT 6

US-09-764-878-36  
; Sequence 36, Application US/09764878  
; Patent No. US20020090615A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA121  
; CURRENT APPLICATION NUMBER: US/09/764,878  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 428  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 287  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-878-36

Query Match 79.0%; Score 15.8; DB 9; Length 287;  
Best Local Similarity 89.5%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTACT 19  
|||||  
Db 102 TGCTAATTCGTGTACT 120

## RESULT 7

US-09-764-860-170  
; Sequence 170, Application US/09764860  
; Patent No. US20020094953A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008  
; CURRENT APPLICATION NUMBER: US/09/764,860  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1198  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 170  
; LENGTH: 287  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-860-170

Query Match 79.0%; Score 15.8; DB 9; Length 287;  
Best Local Similarity 89.5%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTACT 19  
|||||  
Db 102 TGCTAATTCGTGTACT 120

RESULT 8  
US-10-079-854-36  
; Sequence 36, Application US/10079854

; Publication No. US20030054368A1

; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA121C1  
; CURRENT APPLICATION NUMBER: US/10/079,854  
; CURRENT FILING DATE: 2002-02-22  
; Prior application removed - See File Wrapper, or Palm  
; NUMBER OF SEQ ID NOS: 428  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 287  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-079-854-36

Query Match 79.0%; Score 15.8; DB 14; Length 287;  
Best Local Similarity 89.5%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTAATTCGGTGTACT 19  
|||||  
Db 102 TGCTAATTCGTGTACT 120

## RESULT 9

US-10-074-095-170  
; Sequence 170, Application US/10074095  
; Publication No. US2003007704A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008C1  
; CURRENT APPLICATION NUMBER: US/10/074,095  
; CURRENT FILING DATE: 2002-02-14  
; PRIOR APPLICATION NUMBER: 09/764,860  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/235,834  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/234,274  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/234,223  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/228,924  
PRIOR FILING DATE: 2000-08-30  
PRIOR APPLICATION NUMBER: 60/224,518  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,369  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/224,519  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,964  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/241,809  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/249,299  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/236,327  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/241,785  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/244,617  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 60/225,268  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,368  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/251,856  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/251,868  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/229,344  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/234,997  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: 60/229,343  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,345  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,287  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,513  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/223,413  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/229,509  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/236,367  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/237,039  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,038  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/236,370  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/236,802  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,037  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,040  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/240,960  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/239,935  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/239,937  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/241,787  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246,474

PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/246,532  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/249,216  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,210  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/226,681  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225,759  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/225,213  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/227,182  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225,214  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/235,836  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/230,438  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/215,135  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: 60/225,266  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/249,218  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,208  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,213  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,212  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,207  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,245  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,244  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,217  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,211  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,215  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,264  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,214  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,297  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/232,400  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/231,242  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/232,081  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/232,080  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/231,414  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/231,244  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/233,064  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/233,063  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,397  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,399  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/232,401  
PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: 60/241,808  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241,826  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241,786  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/241,221  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246,475  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/231,243  
PRIOR FILING DATE: 2000-09-08

Query Match 79.0%; Score 15.8; DB 14; Length 287;  
Best Local Similarity 89.5%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTATTTGGTGTCTACTG 19  
DB 102 TGTATTTGTGTGTCTACTG 120

RESULT 10  
US-10-128-714-257/C  
Sequence 257, Application US/10128714  
Publication No. US20030119013A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Bo  
APPLICANT: Hu, Weng  
APPLICANT: Tishkoff, Daniel  
APPLICANT: Zamudio, Carlos  
APPLICANT: Erosukhin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
CURRENT FILING DATE: 2002-04-23  
PRIOR APPLICATION NUMBER: US 60/285,697  
PRIOR FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: US 60/287,066  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/295,890  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/303,899  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 60/316,362  
PRIOR FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 257  
LENGTH: 4682  
TYPE: DNA  
ORGANISM: *Aspergillus fumigatus*  
US-10-128-714-257

Query Match 79.0%; Score 15.8; DB 14; Length 4682;  
Best Local Similarity 89.5%; Pred. No. 1.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTAATTCGGTGTCTACTG 20  
DB 4659 GCGAATTCGGCGCTACTG 4641

RESULT 11  
US-10-295-362-19/C  
Sequence 19, Application US/10295362  
Publication No. US20030143745A1  
GENERAL INFORMATION:  
APPLICANT: Aventis Pharma  
APPLICANT: Martinez, Asuncion  
APPLICANT: Kolvek, Steven

TITLE OF INVENTION: Genetically Modified Bacterial Strains and No. US20030143745A1  
TITLE OF INVENTION: Expressing and Assaying Natural Products  
FILE REFERENCE: USAV2001/0145 PCT  
CURRENT APPLICATION NUMBER: US/10/295,362  
CURRENT FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: US 60/335,181  
PRIOR FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19  
LENGTH: 14468  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: pMBD14  
US-10-295-362-19

Query Match 79.0%; Score 15.8; DB 12; Length 14468;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTAATTCGGTGTCTACTG 20  
DB 12215 GCTAATTCGGTGTCTACTG 12197

RESULT 12  
US-10-101-464A-5/C  
Sequence 5, Application US/10101464A  
Publication No. US20030046728A1  
GENERAL INFORMATION:  
APPLICANT: Stradala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgs, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and their use in the Modification of Plant Cell Signaling  
FILE REFERENCE: 11000.1020c2  
CURRENT APPLICATION NUMBER: US/10/101,464A  
CURRENT FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 3340  
TYPE: DNA  
ORGANISM: *Pinus radiata*  
US-10-101-464A-5

Query Match 77.0%; Score 15.4; DB 14; Length 3340;  
Best Local Similarity 94.1%; Pred. No. 2.2e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTAATTCGGTGTCTACTG 18  
DB 1176 GCTAATTCGGTGTCTACTG 1160

RESULT 13  
US-10-044-090-752  
Sequence 752, Application US/10044090  
Publication No. US20020137081A1  
GENERAL INFORMATION:  
APPLICANT: Olga Bandman  
TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
FILE REFERENCE: PA-0028 US  
CURRENT APPLICATION NUMBER: US/10/044,090

```
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 752
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Inocyte ID No. US20020137081A1 034851.1
US-10-044-090-752
```

```
Query Match          76.0%; Score 15.2; DB 13; Length 245;
Best Local Similarity 85.0%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      1 TGCATATTCGCTGCTACCTG 20
          ||||| ||||| |||||
Db      121 TGCATATTCGCTGCTACCTG 140
```

```
RESULT 14
US-10-027-632-35375/c
; Sequence 35375, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 35375
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-35375
```

```
Query Match          76.0%; Score 15.2; DB 13; Length 457;
Best Local Similarity 85.0%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      1 TGCATATTCGCTGCTACCTG 20
          ||||| ||||| |||||
Db      40 TGCATATTCGCTGCTACATG 21
```

```
RESULT 15
US-10-027-632-60413
; Sequence 60413, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 60413
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-60413
```

```
Query Match          76.0%; Score 15.2; DB 13; Length 457;
Best Local Similarity 85.0%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      1 TGCATATTCGCTGCTACCTG 20
          ||||| ||||| |||||
Db      418 TGCATATTCGCTGCTACATG 437.
```

```
Search completed: August 23, 2003, 13:53:17
Job time : 142 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 08:28:23 ; Search time 993 Seconds

(Without alignments)  
823.960 Million cell updates/sec

Title: US-09-927-585A-2

Sequence: 1 tgcataatcggtgctactctg 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl:
1:  gb_da:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_cm:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11:  gb_sts:*
12:  gb_sy:*
13:  gb_un:*
14:  gb_vl:*
15:  em_da:*
16:  em_fun:*
17:  em_hum:*
18:  em_in:*
19:  em_mu:*
20:  em_mu:*
21:  em_or:*
22:  em_ov:*
23:  em_pat:*
24:  em_ph:*
25:  em_pl:*
26:  em_ro:*
27:  em_sts:*
28:  em_un:*
29:  em_vl:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rnd:*
36:  em_htg_man:*
37:  em_htg_vtl:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6 AX301236	AX301236 Sequence
2	20	100.0	20	6 E05232	E05232 Part of DNA
3	20	100.0	20	6 E05256	E05256 Primer for
4	20	100.0	30	6 I34500	I34500 Sequence 9
5	20	100.0	30	6 I57335	I57335 Sequence 9
6	20	100.0	30	6 I73213	I73213 Sequence 9
7	20	100.0	697	14 AF126428	AF126428 Human pap
8	20	100.0	1107	6 AR202651	AR202651 Sequence
9	20	100.0	1128	6 AR202652	AR202652 Sequence
10	20	100.0	1398	6 AR202653	AR202653 Sequence
11	20	100.0	4770	6 AR202671	AR202671 Sequence
12	20	100.0	7902	14 PAPR68	X00203 Genital hum
13	20	100.0	8010	6 AR169144	AR169144 Sequence
14	20	100.0	8010	14 PPR6E	L41216 Human papil
15	20	100.0	8012	14 AF092932	AF092932 Human pap
16	16.8	84.0	3891	14 HP040822	U40822 Human papil
17	16.8	84.0	7887	14 AF436130	AF436130 Human pap
18	16.8	84.0	155468	2 AC129311	AC129311 Mus muscu
19	16.8	84.0	173316	2 AC141726	AC141726 Apis mell
20	16.8	84.0	231361	10 AL731823	AL731823 Mouse DNA
21	16.8	84.0	238650	2 AC135943	AC135943 Rattus no
22	16.8	84.0	242958	2 AC112011	AC112011 Rattus no
23	16.8	84.0	245636	2 AC095871	AC095871 Rattus no
24	16.8	84.0	252185	2 AC122980	AC122980 Rattus no
25	16.8	84.0	290670	2 AC121043	AC121043 Rattus no
26	16.8	84.0	348250	1 AP003592	AP003592 Nostoc sp
27	16.4	82.0	168268	10 AL807796	AL807796 Mouse DNA
28	16.4	82.0	298050	1 AL935260	AL935260 Lactobaci
29	16	80.0	147349	2 AP005800	AP005800 Oryza sat
30	16	80.0	192687	2 AC131316	AC131316 Mus muscu
31	16	80.0	198067	2 BX530057	BX530057 Danio rer
32	16	80.0	213001	2 AC118630	AC118630 Mus muscu
33	16	80.0	214621	2 AC124524	AC124524 Mus muscu
34	16	80.0	223389	2 AC108412	AC108412 Mus muscu
35	15.8	79.0	466	5 AF255339	AF255339 Xenopus 1
36	15.8	79.0	667	1 AF246444	AF246444 Acidithio
37	15.8	79.0	1751	7 BPC31ATMP	X57036 Bacterioph
38	15.8	79.0	2065	3 BC044687	BC044687 Xenopus 1
39	15.8	79.0	2443	3 AB017002	AB017002 Dugesia j
40	15.8	79.0	28939	14 HP031784	U31784 Human papil
41	15.8	79.0	28939	9 AL391808	AL391808 Human DNA
42	15.8	79.0	34216	3 U57054	U57054 Caenorhabdi
43	15.8	79.0	34589	9 AC112240	AC112240 Homo sapi
44	15.8	79.0	37482	3 AC024790	AC024790 Caenorhab
45	15.8	79.0	41489	7 BPH6589	AP006589 Bacteriop

## ALIGNMENTS

RESULT 1  
AX301236  
LOCUS  
DEFINITION  
AX301236  
ACCESSION  
VERSION  
AX301236.1  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1  
ARTHOVS  
TITLE  
JOURNAL

AX301236 20 bp DNA linear PAT 30-NOV-2001  
Sequence 5 from Patent WO0185994.  
AX301236.1 GI:17382323  
synthetic construct  
synthetic construct  
artificial sequences.  
Perseu, S., de Montis, A. and Floris, M.M.  
Method and means for identifying hpv virus  
Patent: WO 0185994-A 5 15-NOV-2001;  
Bionalisti Centro Sud S.N.C. Di Perseu Sinbiado EC. (IT)

FEATURES  
source  
Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="primer representative of region E6 of strains of  
HPV virus having low oncogenic hazard"

BASE COUNT  
3 a 5 c 5 g 7 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
1 TGCTAATTCGGTCTACCTG 20

Db 1 TGCTAATTCGGTCTACCTG 20

RESULT 3  
E05256 20 bp DNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

FEATURES  
source  
Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

BASE COUNT  
3 a 5 c 5 g 7 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
1 TGCTAATTCGGTCTACCTG 20

Db 1 TGCTAATTCGGTCTACCTG 20

RESULT 4  
E05232 20 bp DNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1 (bases 1 to 20)  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

BASE COUNT  
3 a 5 c 5 g 7 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
1 TGCTAATTCGGTCTACCTG 20

Db 1 TGCTAATTCGGTCTACCTG 20

RESULT 5  
E05256 20 bp DNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

FEATURES  
source  
Location/Qualifiers  
1 (bases 1 to 30)  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

BASE COUNT  
9 a 7 c 7 g 7 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
29 TGCTAATTCGGTCTACCTG 10

Db 29 TGCTAATTCGGTCTACCTG 10

RESULT 5  
E05256 30 bp DNA linear PAT 07-OCT-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

synthetic construct  
synthetic construct  
artificial sequences.  
1 (bases 1 to 20)  
Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.  
DETECTION OF HUMAN PAPILLOMA VIRUS  
Patent: JP 1993192200-A 26 03-AUG-1993;  
TAKARA SHUZO CO LTD

OS Artificial gene  
PN JP 1993192200-A/26  
PD 03-AUG-1993  
PF 19-AUG-1991 JP 1991230839  
PR 20-AUG-1990 JP 90P 217067  
PI OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI  
FUJINAGA KEI  
PC C12Q1/70.C12Q1/68;  
CC strandedness: Single;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No.

Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

BASE COUNT  
3 a 5 c 5 g 7 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
1 TGCTAATTCGGTCTACCTG 20

Db 1 TGCTAATTCGGTCTACCTG 20

RESULT 4  
I34500/c 30 bp DNA linear PAT 06-FEB-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1 (bases 1 to 30)  
/organism="unknown"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

BASE COUNT  
9 a 7 c 7 g 7 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
29 TGCTAATTCGGTCTACCTG 10

Db 29 TGCTAATTCGGTCTACCTG 10

RESULT 5  
I57335 30 bp DNA linear PAT 07-OCT-1997  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS



```

VERSION      157335.1  GI:2482399
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 30)
AUTHORS     Guldbrande,S.R. and Kenten,J.H.
TITLE       Method for conducting a polymerase chain reaction using an improved
            electrochemiluminescent label
JOURNAL     Patent: US 5610017-A 9 11-MAR-1997;
FEATURES
SOURCE       1. .30
            /organism="unknown"
BASE COUNT   9 a 7 c 7 g 7 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGCTAATTCGGTCTACCTG 20
    |||
Db 29 TGCTAATTCGGTCTACCTG 10

RESULT 6
LOCUS       173213
DEFINITION  Sequence 9 from patent US 5686244.
ACCESSION   173213
VERSION     173213.1  GI:3009352
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE    1 (bases 1 to 30)
AUTHORS     Guldbrande,S.R. and Kenten,J.H.
TITLE       Method for detecting a nucleic acid analyte using an improved
            electrochemiluminescent label
JOURNAL     Patent: US 5686244-A 9 11-NOV-1997;
FEATURES
SOURCE       1. .30
            /organism="unknown"
BASE COUNT   9 a 7 c 7 g 7 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGCTAATTCGGTCTACCTG 20
    |||
Db 29 TGCTAATTCGGTCTACCTG 10

RESULT 7
LOCUS       AF126428
DEFINITION  Human papillomavirus type 6 E6 protein gene, complete cds.
ACCESSION   AF126428
VERSION     AF126428.1  GI:4761532
KEYWORDS
SOURCE      Human papillomavirus type 6
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE    1 (bases 1 to 697)
AUTHORS     Al,W., Tousseint,E. and Roman,A.
TITLE       CCAAT displacement protein binds to and negatively regulates human
            papillomavirus type 6 E6, E7, and E1 promoters
JOURNAL     J. Virol. 73 (5), 4220-4229 (1999)
MEDLINE     99214361
PUBMED      10196318

```

```

REFERENCE    2 (bases 1 to 697)
AUTHORS     Al,W. and Roman,A.
TITLE       Direct Submission
JOURNAL     Submitted (05-FEB-1999) Microbiology and Immunology, Indiana
            University School of Medicine, 635 Barnhill Drive, Indianapolis, IN
            46202-5120, USA
FEATURES
SOURCE       1. .697
            /organism="Human papillomavirus type 6"
            /mol_type="genomic DNA"
            /strain="W50"
            /db_xref="taxon:31552"
            127. .579
            /codon_start=1
            /product="E6 protein"
            /protein_id="AA029398.1"
            /db_xref="GI:4761533"
            /translation="MESANASTSATITDLCRTKPTNLSMHTLQINCVCFKNAATTAET
            SYVKHLKVLFRGGYPYACACCELEFGKINQYRHFDDYAGATVVEETKODILDVLI
            RCVLCHRPLCEVKVHILKRFKINCTWKRGRCLHCWTCKEDMLP"
CDS
BASE COUNT   233 a 144 c 145 g 175 t
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 697;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGCTAATTCGGTCTACCTG 20
    |||
Db 425 TGCTAATTCGGTCTACCTG 444

RESULT 8
LOCUS       AR202651
DEFINITION  Sequence 19 from patent US 6365160.
ACCESSION   AR202651
VERSION     AR202651.1  GI:21498836
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE    1 (bases 1 to 1107)
AUTHORS     Webb,E. Ann., Margelits,M.Brigid., Cox,J.Cooper., Frazer,I.,
            McMillan,N.Alan,John., Williams,M.Philip.,
            Moloney,M.Bridget.,Holland., and Edwards,S.John.
            Papillomavirus polyprotein constructs
            Patent: US 6365160-A 19 02-APR-2002;
            Location/Qualifiers
            1. .1107
            /organism="unknown"
BASE COUNT   336 a 269 c 213 g 289 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1107;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGCTAATTCGGTCTACCTG 20
    |||
Db 299 TGCTAATTCGGTCTACCTG 318

RESULT 9
LOCUS       AR202652
DEFINITION  Sequence 21 from patent US 6365160.
ACCESSION   AR202652
VERSION     AR202652.1  GI:21498838
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.

```

REFERENCE 1 (bases 1 to 1128)  
AUTHORS Webb,E.An., Margets,M.Brigid., Cox,J.Cooper., Frazer,I., McMillan,N.Alan.John., Williams,M.Philip., Moloney,M.Bridget.Holland. and Edwards,S.John.  
TITLE Papillomavirus polypotein constructs  
JOURNAL Patent: US 6365160-A 21 02-APR-2002;  
FEATURES Location/Qualifiers  
source 1..1128  
BASE COUNT 358 a 275 c 240 g 255 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
Db 299 TGCTAATTCGGTCTACCTG 318

RESULT 10  
AR202653 1398 bp DNA linear PAT 20-JUN-2002  
LOCUS  
DEFINITION Sequence 23 from patent US 6365160.  
ACCESSION AR202653  
VERSION AR202653.1 GI:21498839  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1398)  
AUTHORS Webb,E.An., Margets,M.Brigid., Cox,J.Cooper., Frazer,I., McMillan,N.Alan.John., Williams,M.Philip., Moloney,M.Bridget.Holland. and Edwards,S.John.  
TITLE Papillomavirus polypotein constructs  
JOURNAL Patent: US 6365160-A 23 02-APR-2002;  
FEATURES Location/Qualifiers  
source 1..1398  
BASE COUNT 428 a 331 c 283 g 356 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1398;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
Db 299 TGCTAATTCGGTCTACCTG 318

RESULT 11  
AR202671 4770 bp DNA linear PAT 20-JUN-2002  
LOCUS  
DEFINITION Sequence 45 from patent US 6365160.  
ACCESSION AR202671  
VERSION AR202671.1 GI:21498861  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4770)  
AUTHORS Webb,E.An., Margets,M.Brigid., Cox,J.Cooper., Frazer,I., McMillan,N.Alan.John., Williams,M.Philip., Moloney,M.Bridget.Holland. and Edwards,S.John.  
TITLE Papillomavirus polypotein constructs  
JOURNAL Patent: US 6365160-A 45 02-APR-2002;  
FEATURES Location/Qualifiers  
source 1..4770  
BASE COUNT 1597 a 890 c 1071 g 1212 t  
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 4770;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTCTACCTG 20  
|||||  
Db 2357 TGCTAATTCGGTCTACCTG 2376

RESULT 12  
PAPAB6 7902 bp DNA circular VRL 17-JUL-1998  
LOCUS  
DEFINITION Genital human papillomavirus type 6b (HPV6b).  
ACCESSION X00203  
VERSION X00203.1 GI:60955  
KEYWORDS E1 gene; E2 gene; E4 gene; E5a gene; E5b gene; E6 gene; E7 gene; L1 gene; L2 gene; major capsid protein L1; minor capsid protein L2; regulatory protein E2; replication protein E1.  
SOURCE Human papillomavirus type 6b  
ORGANISM Human papillomavirus type 6b  
VIRUSES: dsDNA viruses, no RNA stage; Papillomaviridae; Papillomavirus.

REFERENCE 1  
AUTHORS Schwarz,E., Durst,M., Demankowski,C., Luttermann,O., Zech,R., Wolfspenger,E., Suhai,S. and zur Hausen,H.  
TITLE DNA sequence and genome organization of genital human papillomavirus type 6b  
JOURNAL EMBO J. 2 (12), 2341-2348 (1983)  
MEDLINE 84131949  
PUBMED 6321162

COMMENT There are two types of open reading frames: putative early open reading frames (E) and putative late open reading frames (L); open reading frames L1 and L2 are expected to code for the viral structural polypeptides  
Location/Qualifiers  
source 1..7902  
BASE COUNT 102 a 331 c 283 g 356 t  
ORIGIN

gene  
TATA\_signal  
CDS  
/gene="E6"  
/codon\_start=1  
/product="E6 protein"  
/protein\_id="CAA25018.1"  
/db\_xref="GI:60956"  
/db\_xref="SWISS-PROT:P06462"  
/translation="MESANASATITIOICKTFLSMHTLIQNCCKNLTATYI STRAHKLVPRGTPVACACCEFHGKINDYRFDAQTATYEETKQDIIIDVLI RYCLCHKPLCEVEKXKILTRARFKLNCYWRGLHCWTQMEDMLP"  
530..826  
/gene="E7"  
530..826  
/gene="E7"  
/codon\_start=1  
/product="E7 protein"  
/protein\_id="CAA25019.1"  
/db\_xref="GI:60957"  
/db\_xref="SWISS-PROT:P06464"  
/translation="MHGRHVLKQIVLDQPPVGLCHCYRLVNSSEDEVNCGD SLPKQHFQIVTCCGCGDSNVLVQCTETIDIREVQQLLTGLNIVCPICAPKT"  
832..2781  
/gene="E1"  
832..2781  
/gene="E1"  
/codon\_start=1  
/product="replication protein E1"

[illegible][illegible]

RESULT 14  
PPHE6E 8010 bp DNA circular VRL 08-AUG-1995  
LOCUS Human papillomavirus, type 6a, complete genome.  
DEFINITION L41216  
ACCESSION L41216.1 GI:940299  
VERSION E1 gene; E2 gene; E4 gene; E5 gene; E6 gene; E7 gene; L1 gene; L2  
KEYWORDS  
SOURCE Human papillomavirus type 6a  
ORGANISM Human papillomavirus type 6a  
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
REFERENCE 1 (bases 1 to 8010)  
AUTHORS Hofmann, K.J., Cook, J.C., Joyce, J.G., Brown, D.R., Schultz, L.D.,  
George, H.A., Rosolowsky, M., Fife, K.H. and Jansen, K.U.  
TITLE Sequence determination of human papillomavirus type 6a and assembly  
of virus-like particles in *Saccharomyces cerevisiae*  
JOURNAL Virology 209 (2), 506-518 (1995)  
MEDLINE 95297152  
PubMed 7778283  
COMMENT On Aug 10, 1995 this sequence version replaced gi:790893.  
source text: Human papillomavirus type 6a DNA.  
FEATURES  
source  
1..8010  
/organism="Human papillomavirus type 6a"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:37122"  
103..355  
/gene="E6"  
103..555  
/gene="E6"  
/note="base 31 is first base after a stop codon; position  
103..105 is first start codon; putative"  
/codon\_start=1  
/protein\_id="AA074211.1"  
/db\_xref="GI:790894"  
/translation="MESANASTATTIDOLCKTFNLSMHTLQINCFCNNALTAETIY  
SYAKOLVLEFGRGYPAACACCELEPHCKINQYRAGVATATVEEELKQDILVLI  
RCYLCHKPLCEPEKVKHILTKARFKLNCNTMKRCLHCWTCMEDMLP"  
531..827  
/gene="E7"  
531..827  
/gene="E7"  
/note="base 441 is first base after a stop codon; position  
531..533 is first start codon; putative"  
/codon\_start=1  
/protein\_id="AA074212.1"  
/db\_xref="GI:790895"  
/translation="MGRHVTLKQDVLDPDPVGLHCYEQLVLSSEDEVDEVQGD  
SPLKQHOIYVCCGCGSNVRLVVGCTETDIREVQQLLTGLDLYCPICAKT"  
833..2782  
/gene="E1"  
833..2782  
/gene="E1"  
/note="base 716 is first base after a stop codon; position  
833..835 is first start codon; putative"  
/codon\_start=1  
/protein\_id="AA074213.1"  
/db\_xref="GI:790896"  
/translation="MADSGTNEGSGCGMVEVAIVQPTQIISDDEBEVEDSG  
YDMVDFDSNTHNSLEAOLFNROEADTHATVODLKKRYLGSPPVSPITIAEAV  
ESEISPRIDALITLQPKVKRRLLQTRBLDTSYGYSVEAGTGTQVEKMGVPENG  
DQEDKIDRGEHEETAEAPNSVREAGTAGLLELCKDLRAALLGKKEEGL  
SFIDILRPFKSDKTCADWVAVGFIHHSISAPQILIEPLSLVAHIQMLTKAMAVL  
LVAFKTKSRSTVARTLATLINTPDNOMLIEPRKIOSGVALAWFERGISNASTVI  
GAPWITRQVIEHGLADSOBKLEFMOMAYDNDICSESEIAFEYAGRGDDSNARA  
PLNSMOKAYVADCATMCRHRTKHAERKMSIKQWIKHRSKTEGTGMKPIYQFLRHQ

NIETPLSKRLMLHGTPKKNCIAIYGPPTGKSYNCMSLISLGGYVISHVNSSH  
FWLOPLVDKVALLDADATOPCWIYMDPITMRLLDGNPMSIDRKHALTLKCPPLVAT  
SNIDITKEEKYKYLHTRVTFPPNPPFBNNGNAVYELSNANKCEFERLSSLDIQ  
DSEDEEDSGNSOAFRCVGVVRTL"  
2724..3830  
/gene="E2"  
2724..3830  
/gene="E2"  
/note="base 2697 is first base after a stop codon;  
position 2724..2726 is first start codon.; putative"  
/codon\_start=1  
/protein\_id="AA074214.1"  
/db\_xref="GI:790897"  
/translation="MEAIKRLDACQEBLLLEYENSTDLKRVHLHKCMHRESVLLY  
KAKQGLSHIMQVAVPLKVSARGHNAIEMQMLDESLLKTEYSHEPTLQETSEM  
QTPKRCFKRKGKTYVEKFCANNTMDYVVDVYQDTSWVKVSHVMAKAGIYLT  
CGOFTYVNVNKEAEKYSTKQWECVGVISPVASSTQGVESIPETSTPAQ  
TSTPVSSSTQEDAVOTPPRKARAGVOQSCNALCYAHIGPDSGNHNLITNNHOHR  
BNNSSSATPIVQOGESNCLCFRYRLNDKRHLFDLISSTWMAHSPKAPKHAIVT  
VTYHSEBQROQPLNVYKIPITIRKLGMSLHL"  
3286..3585  
/gene="E4"  
3286..3585  
/gene="E4"  
/note="base 3241 is first base after a stop codon;  
position 3286..3288 is first start codon; putative"  
/codon\_start=1  
/protein\_id="AA074215.1"  
/db\_xref="GI:790898"  
/translation="MAAOLVYLHLHYLALHKKYPLNLHPHPRPPLCQAPRKQ  
CKRRLNEHESNSHLARPCWPLTDPVETFTSSLTITSTKGTIVVQVLRH"  
3888..4163  
/gene="E5"  
3888..4163  
/gene="E5"  
/codon\_start=1  
/protein\_id="AA074216.1"  
/db\_xref="GI:790899"  
/translation="MEVYPVQIAGTSTLILPVIARVGVFSILLTWISDFIVYT  
SVLVTLTLLMLLTLTLPLOFLTLVLCVCPALYIHHYVNTQQ"  
4424..5803  
/gene="L2"  
4424..5803  
/gene="L2"  
/note="base 4379 is first base after a stop codon;  
position 4424..4426 is first start codon; putative"  
/codon\_start=1  
/protein\_id="AA074217.1"  
/db\_xref="GI:790900"  
/translation="MAHSRARRRRKASATOLYQCKLGTGCPDVIYKVEHNTADQI  
LKMGLGVFPGIGTGTGSCGRTGCVPLGCSAKPSITSGPMARPVYVPEVAPSDP  
STVSLIESAIIINAGAPETVPPAHGFPITTSSETTTPAILDVSATSHTTTIFRNPV  
TEPSTVQOPPEVANGHLLISAFTTTSPIETIPDITVYISSDSGPPSPVPEGTAP  
RPRGLVSRALHQVQVTDPAFLSTPQRLITVDNPPYEBEDVSVQSHSHINADEAR  
MDIILHRPRAHLSRGLVAFVSRIGQSGMHRSGKHIGARLHYFEDISPIQAAEIE  
MHPIVAAQDDPFDIYASPEPDINTQHPVNIISPTLYTSPNVTQWGMGTVPLSI  
PNDFPLQSGPDTTPPTAPMGTPFSPVYALPGLGVITGSGFYLHPAYIFAKRRKRI  
PLFESDVAA"  
5790..7292  
/gene="L1"  
5790..7292  
/gene="L1"  
/note="base 5679 is first base after a stop codon;  
position 5790..5792 is first start codon.; putative"  
/codon\_start=1  
/protein\_id="AA074218.1"  
/db\_xref="GI:790901"  
/translation="MWRPSDSYVPPPNPVSUYATAYVTRTNIFFHASSRLLAY  
GHYPSILRANKTVKPVKSGYQYREKVVLDPPNKFALPDSLSLEPTQRLVWACTGL  
EYGRGQPIKRGVSGHPFLNKDYDVSSGSGGPNQDNVNVGMDEKQQLQVACATGL  
LGEHMGKQKQCNTPVQAGDCEPLTISYIODGMVDTFGAMDFADQNLNKSDVP  
DICGTTCCVPTDQVMAADPYGRLEFFPKRDMFMRHFPNRAEGVEGEVLPDILIKGS  
GNRTVSGSIYVNTPSGSLVSSEADLPFKPIYTLQAKQGHNNKICGNQDLFTYVDTTR

STNMTCASATTSSTYNTSDYKRYMRHVEEYDIOFLQCLSTLSAEWMAYIHTMNP  
VLEDMNGLSPSPNGTLEDYRYVQSOAITCQPTPEKEKPDYKNI,SEWEVILKEKEF  
SEBDDYPLCRKLLQSGYGRSSIRGTGYRPAVSKASAPKRRKRAKTKR"

BASE COUNT 2470 a 1527 c 1718 g 2295 t

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 8010;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGGTAATTCGCTACCTG 20  
|||||  
401 TGGTAATTCGCTACCTG 420

RESULT 15  
AF092932 8012 bp DNA linear VRL 14-NOV-2001  
LOCUS Human papillomavirus type 6, complete genome.  
DEFINITION AF092932  
VERSION AF092932.1 GI:6002612  
KEYWORDS  
SOURCE Human papillomavirus type 6  
ORGANISM Human papillomavirus type 6  
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;  
Papillomavirus.  
REFERENCE 1 (bases 1 to 8012)  
AUTHORS Kovelman, R., Bilter, G. K., Roman, A., Brown, D. R. and Barbosa, M. S.  
TITLE Human papillomavirus type 6: classification of clinical isolates  
JOURNAL J Gen. Virol. 80 (Pt 9), 2445-2451 (1999)  
MEDLINE 99429608  
PUBMED 10501500  
REFERENCE 2 (bases 1 to 8012)  
AUTHORS Kovelman, R., Bilter, G. K., Roman, A., Brown, D. R. and Barbosa, M. S.  
TITLE Direct Submission  
JOURNAL Submitted (17-SEP-1998) Virology, Signal Pharmaceuticals, Inc.,  
5555 Oberlin Drive, San Diego, CA 92121, USA  
location/Qualifiers  
1. 8012  
/organism="Human papillomavirus type 6"  
/mol\_type="genomic DNA"  
/strain="Human papillomavirus type 6vc"  
/db\_xref="taxon:31352"  
103. 555  
/gene="E6"  
103. 555  
/gene="E6"  
/codon\_start=1  
/product="E6"  
/protein\_id="AA00064.1"  
/db\_xref="GI:6002618"  
/translation="MESANASTNTTIDOLCKTFENLSMHTIQINCVCKNAITTAETI  
SYVKOLYVFGYRGYACACLEFGKRIQVRYHDPYAGYATVEETKODIIDLVI  
RCYLCHKPLCEVKKHILTKAFIKLNTCKWGRCILCHMTTCEMDLP"  
531. 827  
/gene="E7"  
531. 827  
/gene="E7"  
/codon\_start=1  
/product="E7"  
/protein\_id="AA00065.1"  
/db\_xref="GI:6002619"  
/translation="MHGRHYTLKDIYDLQPPDPVGLHCYEQLVDSSEDEVDEYGD  
SPLKHQHYIVTCCGCDSDSVRLVYQCTERDIREYQQLLGITINIVCPICAPKT"  
833. 2782  
/gene="E1"  
833. 2782  
/gene="E1"  
/codon\_start=1  
/product="E1"  
/protein\_id="AA00059.1"  
/db\_xref="GI:6002613"

/translation="MADDSGTENEGSGCTGFWNEAIVQHPCTQISDDEDEVEDSG  
YDMVDFIDSDINITHNSLEAQLFNRQADHYATVODLKRKYLGSPVSPINTIAYAV  
ESEISIPDLAIKILTRQPKYKRLPQTRLELDSGYGSEAEAGTQVQVGHVPEENG  
DGOEKDGRDIEGEHTEAEAPNSVREHAGTGIILELCKLRAALLGKPFCEGL  
SFIDLRPEKSDKTCADMYAVAGGJIHHSISEAFOLIEPLISYAHIOITNMAGVL  
LVLLRKNKSRSTVARTATLNLINPENQMLIEPPKIQSGVALYVFRGISNASTVI  
GEAPENITQTVIEHGLADQFRLTEVQVAYNDICESEIAFEYAQRQDPSNARA  
FLNSNQAKYKDCATMCRHYKHAEMRKMSIKQMIKRGSKIGTGNMKEPIVQLRQ  
NIFIEPLSKFELMLGTCPKNCIAIYGPDPDCTSYCKMSLISFLGTQVISHVSSSH  
FMLOPLVDALVALLDQATPCWIMDYMNKLIDGPNMSIDRKHALTLTIKCPPLVT  
SNIDIKKEEKYKYLHTRYTFEPFPNPPEDRONNAYELSNAMWCKFEPLSSLDIO  
DSEDEEDGNSQAFRCVPGTVRTL.  
2724. 3830  
/gene="E2"  
2724. 3830  
/gene="E2"  
/codon\_start=1  
/product="E2"  
/protein\_id="AA00060.1"  
/db\_xref="GI:6002614"  
/translation="MEAIARKDACOEQLLEYENSTDLNKHVLIHMKCIHRSVLLY  
KAKOMGLSHIGMOVVPLPKYSEAKGNHAIEMOKHLELTKETYSMEPMTLOETSEYM  
OTPKRCFCRGRGTVEYKFDGCANMTMDYVMTDYVDYODDSWKYHSDAKGIYTT  
CGPRTYTVNPKVEAEKYSTKQWECYGTVCSPASVSTTOEVSIPSTYTRQ  
TSTPVSSSTQEDAVQTPPKRARGVQDSFNALCAHIGVDGNNHVTINNDQHOR  
RNNNSNSATPIVOFGESNCLKCFRRLNHRHLDLISSTWMASSKAPKHAATV  
VYHSEORQOQFLNVVYIPIPIRHLKLFMSHLL"  
3256. 3585  
/gene="E4"  
3256. 3585  
/gene="E4"  
/codon\_start=1  
/product="E4"  
/protein\_id="AA00061.1"  
/db\_xref="GI:6002615"  
/translation="MGAPNNGKYVMAQALVILHLYALHKKYPLNLHTPPHRRP  
LCQAPRKTCCKRLENEHESHSLATPCVWPTLDPWYETTTSSLTITTSKREGT  
VTVQLRL"  
3888. 4163  
/gene="E5A"  
3888. 4163  
/gene="E5A"  
/codon\_start=1  
/product="E5A"  
/protein\_id="AA00062.1"  
/db\_xref="GI:6002616"  
/translation="MEVVPVOIAGTSTLILPVIILFVVCFSIILIVMSDFIYVT  
SVVLTLLYLLMLLTLPLOFFLLTLVLCYCPCACIHHYIVNQO"  
4160. 4378  
/gene="E5B"  
4160. 4378  
/gene="E5B"  
/codon\_start=1  
/product="E5B"  
/protein\_id="AA00063.1"  
/db\_xref="GI:6002617"  
/translation="MMLTQFNDGDTWGLMLCAFLVGLLMLHMYRAVQDCKHTK  
CNCKNKNCDNDVYTHMYTAGDIYVNN"  
4424. 5803  
/gene="L2"  
4424. 5803  
/gene="L2"  
/codon\_start=1  
/product="L2"  
/protein\_id="AA00067.1"  
/db\_xref="GI:6002621"  
/translation="MAHSRAKRRKRASATOLYOTCKITGTCPPDVIPKVENHTADQI  
LKWGSLVFFGGIGIGSGTGTGTVPLGTSAKPSITSGPARPVPVVEVPAPSDP  
SIVLSIEESAIINAGAPEIYPPAHGFTIISSETTPAILDVSTSTTYSIRNPVF  
TEPSVIOPOPVEANGAIIISAPITISHPIEELPLDFTVSSDSGPTSSTPVGPAP  
RPRVGLYSRALHOVOYDTPALSTPORLITYDNDYEGEDVSQFISHNADDEAF  
MDIIRLRPAIASRGLYRSRIGQSRGSHMTSGKHIGARIIHFYDPSIAQAAELE  
MHPLVAAQEDTFDIASEFEPDINPTQHPVTNISDTYLTSTPMTVVOPMGNTVPLSI

gene  
PNDLFQSGPDITFEPTAPMGTFPSFVTPALPTGPVEITGSGFYLHPAMYFAKRRKRI  
PLEFSDVNA"  
5790..7292  
/gene="L1"  
5790..7292  
/gene="L1"  
/codon\_start=1  
/product="L1"  
/protein\_id="AAF00066.1"  
/db\_xref="GI:6002620"  
/translation="MMRSDSTVYVPPNPVSKVYATDYYTRNIFYHASSRLAV  
GHPYSIRANKTVYPKYSGYQYRFKYLDPNKFALPDSLEPPTQRLVMACGL  
EVRGQPLGCVSGHPLNKYDDVENSNGGNGODNRVNGMDYKQQLCNVGCAP  
LSEHWKKGKQCTNTPVQAGDCPELITSVIODGMVDIGFAMNFADLQTNKSDVPI  
YICGTCRYPDYLMADPYGDRLEFLRKEQMFARHFFNRAGEVGEVPDYLIIKGS  
GNRTSVSSSIYVNPSSLSVSEAOLEFNKPYLQKAGCHNNGICGNOLFVTVDDTR  
STNMTLCASVTTSSTYTNSDYKEYMRHYEYDLOFIFOLCSITLSAEYMAVYHTMNP  
VLEDMNGLSPPNCTLEDYTRYVQSQAITCKPTPEKEKPDPTKNSFWENLKEKE  
SSELQYPLGRKFLQSGIRGSSIRTVGKRPAVSKASAPAKRRAKTKR"

BASE COUNT 2478 a 1526 c 1713 g 2293 t  
ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 8012;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTAATTCGGTGTCTACCTG 20  
|||||  
DB 401 TGCTAATTCGGTGTCTACCTG 420

Search completed: August 23, 2003, 11:11:37  
Job time : 997 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:14:58 ; Search time 1848 Seconds  
(without alignments)  
263.035 Million cell updates/sec

Title: US-09-927-585a-1

Perfect score: 20  
Sequence: 1 tgcataaacctgtgtcc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_png:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	90.0	670	12	BM076659 TREST-A05
2	17.4	87.0	629	28	AQ291095 nbxb0038C
3	17	85.0	831	28	AZ081985 UP_309-16
4	17	85.0	863	29	CNS009PI AL054375 Drosophila

	5	16.8	84.0	297	9	AM454500
C	6	16.8	84.0	308	14	CA620055
C	7	16.8	84.0	370	12	BM103817
C	8	16.8	84.0	410	10	BE760467
C	9	16.8	84.0	425	12	BM201540
C	10	16.8	84.0	473	28	AQ228742
C	11	16.8	84.0	508	9	A1443887
C	12	16.8	84.0	645	14	CA645372
C	13	16.8	84.0	659	12	B1830009
C	14	16.8	84.0	875	13	B0941677
C	15	16.8	84.0	929	10	BG335280
C	16	16.8	84.0	999	13	BX329000
C	17	16.4	82.0	322	10	BG405554
C	18	16.4	82.0	328	10	BG405573
C	19	16.4	82.0	354	10	BG406469
C	20	16.4	82.0	374	9	AL817342
C	21	16.4	82.0	388	10	BG653697
C	22	16.4	82.0	413	10	BG652756
C	23	16.4	82.0	429	10	BG363250
C	24	16.4	82.0	430	10	BG363563
C	25	16.4	82.0	431	10	BG652072
C	26	16.4	82.0	442	10	BG363641
C	27	16.4	82.0	454	10	BG651780
C	28	16.4	82.0	461	10	BE803096
C	29	16.4	82.0	476	10	BG363283
C	30	16.4	82.0	479	10	BG363409
C	31	16.4	82.0	482	10	BG725786
C	32	16.4	82.0	491	10	BG363306
C	33	16.4	82.0	492	10	BG363341
C	34	16.4	82.0	492	10	BG652372
C	35	16.4	82.0	493	10	BG359585
C	36	16.4	82.0	497	10	BG725368
C	37	16.4	82.0	499	10	BG651803
C	38	16.4	82.0	499	10	BG653953
C	39	16.4	82.0	502	10	BG651372
C	40	16.4	82.0	503	10	BG725471
C	41	16.4	82.0	504	10	BF423738
C	42	16.4	82.0	505	10	BG363449
C	43	16.4	82.0	508	10	BG406572
C	44	16.4	82.0	508	10	BG651997
C	45	16.4	82.0	511	12	BG789864

## ALIGNMENTS

RESULT 1  
BM076659 670 bp mRNA linear EST 05-FEB-2002  
TREST-A0592 TREST-A Hypocrea jecorina cDNA clone Tr-A0592 5', mRNA  
sequence.  
ACCESSION BM076659  
VERSION BM076659.1 GI:18497841  
KEYWORDS  
SOURCE EST.  
ORGANISM Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Hypocreaeae; Hypocrea.  
REFERENCE  
1 (bases 1 to 670)  
Chamberg, F.S., Bonaccorsi, E.D., Ferreira, A.J.S., Ramos, A.S.P.,  
Ferreira, J.R.Jr., Abrahao-Neto, J., Farah, J.P.S. and El-Dorri, H.  
Elucidation of the metabolic fate of glucose in the filamentous  
fungus *Trichoderma reesei* using expressed sequence tag (EST)  
analysis and cDNA microarrays  
J. Biol. Chem. 277 (16), 13963-13968 (2002)  
JOURNAL MEDLINE  
PUBMED  
COMMENT  
Contact: El-Dorri, Hamza  
Department of Biochemistry  
Institute of Chemistry, University of Sao Paulo  
Avenida Professor Lineu Prestes, 748, Sao Paulo, SP, 05508-900,  
BRASIL  
Tel: (55) 11-38183848

Fax: (55) 11-38183848  
Email: dorry@iq.usp.br  
PCR Primers  
FORWARD: Universal M13 forward primer  
BACKWARD: Universal M13 reverse primer  
Plate: 7 row: B Column: 4  
Seq primer: M13 reverse primer  
High quality sequence stop: 670  
POLYA-No.

#### FEATURES

source

Location/Qualifiers  
1..670  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="QM9414 (ATCC26921)"  
/db\_xref="taxon:51453"  
/clone="Tr-A0592"  
/sex="Asexual"  
/tissue\_type="Mycelia"  
/dev\_stage="18 hr glycerol-grown culture"  
/lab\_host="E. coli SDR cells (kanamycin resistant)"  
/clone\_lib="TREST-A"  
/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2: XhoI; anamorph-Trichoderma reesei; Cloned unidirectionally, 5' end of the cDNA cloned into EcoRI site of pBluescript. Primer: Oligo (dT). Average insert size: 1,2 kb; Uni-ZAP XR vector system -5' adaptor sequence: 5'GAATCGGCACGAG3'-3' adaptor sequence: 5'CTCGAGTTTTTTTTTTTTTTT3'."

BASE COUNT  
164 a 184 c 135 g 185 t 2 others

ORIGIN

Query Match  
Best Local Similarity 90.0%; Score 18; DB 12; Length 670;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTGTGT 18  
|||||  
Db 425 TGTCAAAAACCGTGTGT 442

RESULT 2  
AO291095/c 629 bp DNA linear GSS 03-DEC-1998  
LOCUS  
DEFINITION nbx0038C14r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbx0038C14r, genomic survey sequence.

ACCESSION AO291095.1 GI:3952385  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 629)  
Wing,R.A. and Dean,R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: GGAACACGATGACCATC  
Class: BAC ends  
High quality sequence stop: 481.  
Location/Qualifiers

FEATURES  
source  
1..629  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/strain="japonica"

/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="nbx0038C14r"  
/tissue\_type="leaf"  
/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice BAC library"  
/note="Vector: pBelOAC11; Site\_1: HindIII; Site\_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT  
223 a 134 c 109 g 162 t 1 others

ORIGIN

Query Match  
Best Local Similarity 87.0%; Score 17.4; DB 28; Length 629;  
Best Local Similarity 94.7%; Pred. No. 2.1e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGTCAAAAACCGTGTGT 19  
|||||  
Db 262 TTTCAAAAACCGTGTGT 244

RESULT 3  
AZ081985/c 831 bp DNA linear GSS 28-JAN-2002  
LOCUS  
DEFINITION UP\_309-16G.SP6 RPII Human Male BAC Library Homo sapiens genomic clone 309-16G, genomic survey sequence.

ACCESSION AZ081985.2 GI:18379387  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 831)  
Cheung,V.G., Dalrymple,H.L., Narasimhan,S., Watts,J., Schuler,G., Raap,A.K., Morley,M. and Bruzel,A.  
A resource of mapped human bacterial artificial chromosome clones  
Genome Res. 9 (10), 989-993 (1999)  
99455100  
10523527  
On Jan 28, 2002 this sequence version replaced gi:7709240.  
Contact: Arcaro MA, Morley M, Burdick J, Cheung VG  
Department of Pediatrics  
University of Pennsylvania  
3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA  
Tel: 215 590 2664  
Fax: 215 590 3709  
Email: mlenox@mail.med.upenn.edu  
Plate: 309 row: G column: 16  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers

FEATURES  
source  
1..831  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="309-16G"



```

/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPC11 Human Male BAC Library"
/notes="Vector: pBAC3.6; RPC11 Human Male BAC Library"
BASE COUNT      226 a      189 c      174 g      242 t
ORIGIN

Query Match      85.0%; Score 17; DB 28; Length 831;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GTCAAAACCGTGTGT 18
        |||
Db      655 GTCAAAACCGTGTGT 639

RESULT 4
CNS009PI
LOCUS
DEFINITION
  CNS009PI      863 bp      DNA      linear      GSS 03-JUN-1999
  Drosophila melanogaster genome survey sequence FET3 end of BAC #
  BACR0D02 of RPC1-98 library from Drosophila melanogaster (fruit
  fly); genomic survey sequence.
ACCESSION
  AL054375      GI:4935346
VERSION
  AL054375.1    GI:4935346
KEYWORDS
  Drosophila melanogaster (fruit fly)
SOURCE
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 863)
REFERENCE
  Genoscope.
  Direct Submission
  Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazutoyo Osoegawa and
  Aaron Mammeser in Pleter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPC1-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain y2; cn bw sp; the same strain used for the BDGP's
  P1 and EST libraries. A more detailed description of the library
  and how to order individual BAC clones, the entire library, or
  filters for hybridization from the BACPAC Resource Center can be
  found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
  location/Qualifiers
FEATURES
  source
    1..863
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone_lib="BACR0D02"
    /clone_1lb="RPC1-98"
    /note="end : FET3"
BASE COUNT      469 a      155 c      91 g      77 t      71 others
ORIGIN

Query Match      85.0%; Score 17; DB 29; Length 863;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 TGTCAAAACCGTGTGTTC 19
        |||
Db      560 TGTCAAAACCGTGTGTTC 578

RESULT 5
AM454500

```

```

LOCUS
DEFINITION
  AM454500      297 bp      mRNA      linear      EST 17-FEB-2000
  zebrafish Embryonic Heart CDNA library Danio rerio CDNA
  5', mRNA sequence.
ACCESSION
  AM454500
VERSION
  AM454500.1    GI:6995287
KEYWORDS
  EST.
SOURCE
  Danio rerio (zebrafish)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
  Cyprinidae; Danio.
  1 (bases 1 to 297)
REFERENCE
  Ton,C., Mably,D.D., Dempsey,A.A., Hwang,D.M., Fishman,M.C. and Liew
  ,C.C.
  Identification and Characterization of Expressed Sequence Tags from
  an Embryonic Zebrafish Heart cDNA Library
  Unpublished
  Contact: Liew CC
  Brigham and Women's Hospital
  Harvard Medical School
  75 Francis St. Boston, MA 02115, USA
  Tel: 6177328915
  Fax: 6179750995
  Email: cliw@rics.bwh.harvard.edu
PCR Primers
  FORWARD: 5' GCCAGCTCGAATTACCTCCTCAATAAGG 3'
  BACKWARD: 5' CCAGTGATTTGTAATGCGCTCCTACTTATAGGG 3'
  Seq primer: 5' GAATTATACCTCTCACTAAGG 3'.
  Location/Qualifiers
FEATURES
  source
    1..297
    /organism="Danio rerio"
    /mol_type="mRNA"
    /dev_stage="embryonic day 3 post-fertilization"
    /db_xref="taxon:7955"
    /lab_host="E.coli XL1-Blue mrf"
    /clone_1lb="Zebrafish Embryonic Heart CDNA Library"
    /note="Organ: heart; Vector: lambda ZAP Express; Site:1:
    EcoRI; Site:2: XhoI; mRNA was purified from embryonic
    zebrafish hearts (3 day post-fertilization). cDNA was
    synthesized using a XhoI-oligo dT adaptor-primer. EcoRI
    adaptors were ligated, followed by digestion with XhoI,
    for directional cloning into pre-digested lambda ZAP
    Express vector."
BASE COUNT      101 a      56 c      61 g      79 t
ORIGIN

Query Match      84.0%; Score 16.8; DB 9; Length 297;
Best Local Similarity 90.0%; Pred. No. 3.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TGTCAAAACCGTGTGTTC 20
        |||
Db      107 TGTCAAAACCGTGTGTTC 126

RESULT 6
CA620055/c
LOCUS
DEFINITION
  CA620055      308 bp      mRNA      linear      EST 23-NOV-2002
  w1ln.pk0052.f8 w1ln Triticum aestivum CDNA clone w1ln.pk0052.f8 5'
  end, mRNA sequence.
ACCESSION
  CA620055
VERSION
  CA620055.1    GI:25198351
KEYWORDS
  EST.
SOURCE
  Triticum aestivum (bread wheat)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
  Triticeae; Triticum.
  1 (bases 1 to 308)
REFERENCE
  Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
  Miao,G., Caraher,N. and Hanafey,M.K.
  Dupont Wheat cDNA Sequence
  Unpublished

```

COMMENT Contact: Scott V. Tingey  
Crop Genetics  
E. I. Dupont de Nemours and Company  
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA  
Tel: 302-631-2602  
Fax: 302-631-2607  
Email: Scott.V.Tingey@usa.dupont.com  
Seq primer: M13.

FEATURES  
source Location/Qualifiers  
1..308  
/organism="Triticum aestivum"  
/db\_xref="taxon:4565"  
/clone\_wln="PK052.16"  
/rname="leaf"  
/clone\_lib="wln"  
/note="Vector: pBluescript SK+; Site.1: EcoRI; Site.2: XhoI; Wheat (Triticum aestivum L.) leaf 7 day old seedling  
light grown (normalized)"

BASE COUNT 94 a 69 c 44 g 101 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 308;  
Best Local Similarity 90.0%; Pred. No. 3.9e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTCAAAAACCGTGTGTC 20  
|||||  
30 TGTCAAAAACCATGTGTC 11

DB

RESULT 7  
BM103817 370 bp mRNA linear EST 21-NOV-2001  
f39c12.x1 Sugano SJD adult male Dario rerio cDNA clone  
IMAGE:5410942.3' similar to SW:RPAS\_HUMAN\_015160 DNA-DIRECTED RNA  
POLYMERASE I 40 KD POLYPEPTIDE ;, mRNA sequence.

LOCUS BM103817  
DEFINITION

ACCESSION BM103817 GI:17034887  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished  
Other\_ESTs: f39c12.y1  
Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbraflsh@wustl.edu  
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA  
Sequencing by: Washington University Genome Sequencing Center Clone  
distribution information can be found through the I.M.A.G.E.  
Consortium/ILNU, send email to: info@image.llnl.gov  
Trace considered overall poor quality  
High quality sequence stop: 1.

FEATURES  
source Location/Qualifiers  
1..370  
/organism="Dario rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:5410942"

/sex="male"  
/rname="whole body"  
/dev\_stage="adult"  
/lab\_host="DH10B (phage resistant)"  
/clone\_lib="Sugano SJD adult male"  
/note="Vector: pME185-FU3; Site.1: DraIII (CACCATG);  
Site.2: DraIII (CACATG); 1st strand cDNA was primed  
with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT];  
double-stranded cDNA was ligated to a DraIII adaptor  
[GTGTGGCTTCTG], digested and cloned into distinct DraIII  
sites of the pME185-FU3 vector (5' site CACATG, 3' site  
CACCATG). XhoI should be used to isolate the cDNA  
insert. Site selection was performed to exclude fragments  
<1.5kb. Library constructed and donated by Dr. Sumio  
Sugano (University of Tokyo Institute of Medical Science).  
Custom primers for sequencing: 5' end primer  
CTTGTGCTTAAGCTGCG and 3' end primer  
CGACCTGCAGCTCGACACA."

BASE COUNT 93 a 114 c 81 g 82 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 370;  
Best Local Similarity 90.0%; Pred. No. 4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTCAAAAACCGTGTGTC 20  
|||||  
200 TGTCAAAAACCGTGGGCC 181

DB

RESULT 8  
BE760467 410 bp mRNA linear EST 24-OCT-2000  
an.1726 Aspergillus niger, pYES2 (XhoI-EcoRI) Aspergillus niger  
cDNA clone 1726 3', mRNA sequence.

LOCUS BE760467  
DEFINITION

ACCESSION BE760467 GI:10183104  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished  
Contact: Tsang, A.  
Dept. of Biology  
Concordia University  
1455 deMaisonneuve Blvd. West, Montreal, Quebec, CANADA, H3G 1M8  
Tel: 514-848-3405  
Fax: 514-848-4504  
Email: tsang@vax2.concordia.ca  
PCR Primers  
BACKWARD: GCGCTGAATGTAACGTCGACATAC  
Insert Length: 410 Std Error: 0.00  
POLYA-No.

FEATURES  
source Location/Qualifiers  
1..410  
/organism="Aspergillus niger"  
/mol\_type="mRNA"  
/db\_xref="taxon:5061"  
/clone="1726"  
/lab\_host="E. coli"  
/note="Vector: pYES2; Site.1: Xho-I; Site.2: EcoRI; cDNA  
was synthesized with ZAP kit (Stratagene) using poly(A)+  
RNA isolated from Aspergillus niger grown in both complete  
and minimal media. Synthesis was primed with oligo(dT)  
primer/XhoI-linker. EcoRI adaptors were later ligated to  
EcoRI-XhoI-digested pYES2 (Invitrogen Corp). This vector  
permits expression of our library in yeast."

BASE COUNT	103 a	116 c	101 g	81 t	9 others
Query Match					
Best Local Similarity	84.0%	Score 16.8;	DB 10;	Length 410;	
Matches	18;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
Db	270	TATCAAAAACAGTTGTCTC	251		
RESULT 9					
LOCUS	BM201540				
DEFINITION	BM201540				
ACCESSION	BM201540				
VERSION	BM201540.1				
KEYWORDS	EST.				
SOURCE	Mus musculus				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Mus.				
AUTHORS	1 (bases 1 to 425)				
TITLE	Piao,X., Karyul,G.J., Dudekula,D.B., Qian,X., Tanaka,T., Lim,M.K., Luo,A. and Ko,M.S.H.				
JOURNAL	Systematic Analyses of NIA Mouse 7.5-dpc Whole Embryo cDNA Library (Long)				
COMMENT	Unpublished				
	Contact: Dawood B. Dudekula				
	Laboratory of Genetics				
	National Institute on Aging/National Institutes of Health				
	333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA				
	Email: cdna@gsun.grc.nia.nih.gov				
	Plate: C0227 row: A column: 12				
	Seq primer: -21M13 Forward				
	High quality sequence stop: 425				
	POLYA-Yes.				
FEATURES					
source	Location/Qualifiers				
	1..425				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/strain="C57BL/6J"				
	/db_xref="niaEST:C0227A12-3"				
	/db_xref="taxon:10090"				
	/clone="C0227A12"				
	/tissue_type="whole embryo including extraembryonic				
	tissues at 7.5-days postcoitum"				
	/dev_stage="7.5-days postcoitum"				
	/lab_host="DH10B"				
	/clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library				
	(Long)"				
	/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:				
	NotI; Mouse cDNA project by the Laboratory of Genetics,				
	National Institute on Aging (NIA), Intramural Research				
	Program, NIH ( <a href="http://1gsun.grc.nia.nih.gov/cDNA">http://1gsun.grc.nia.nih.gov/cDNA</a> ). This is				
	a long-transcript enriched cDNA library (Ref. Genome Res.				
	11: 1553-1558 (2001). [PMID: 11544199]) . Total RNAs were				
	extracted from a pool of four embryos at 7.5-days				
	postcoitum. Double-stranded cDNAs were synthesized with an				
	Oligo(dT) primer (Invitrogen:				
	5'-pGACATCTTCAGATCCGACGCGCCGCTTTTTTTTTTTT-3') from				
	7 ug of total RNA, treated with T4 DNA polymerase, and				
	purified by ethanol-precipitation. The cDNAs were ligated				
	to lone-linker L1-SalI, purified by phenol/chloroform, and				
	separated from free linkers by Centricon 100. Then, the				
	cDNAs were amplified by long-range high fidelity PCR using				
	Ex Taq polymerase (Takara) with a primer SalI-S. The				
	products were purified by phenol/chloroform and Centricon				
	100. The cDNAs were digested with SalI and NotI enzymes				
	and cloned into SalI/NotI site of pSPORT1 plasmid vector.				
	The DH10B E. coli host was transformed with the ligation				

```

mixture by the standard chemical method. The average
insert size is about 2.2 kb. The library was constructed
by Yulan Piao (NIA)."
```

BASE COUNT	144 a	87 c	76 g	118 t
ORIGIN				
Query Match	84.0%;	Score 16.8;	DB 12;	Length 425;
Best Local Similarity	90.0%;	Pred. No. 4.1e+02;		
Matches	18;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
OY	1	TGTCAAAACCGTTGTGTC	20	
DB	234	TGTCAAAACCGTTGTGTC	253	
RESULT 10				
LOCUS	AO228742	473 bp	DNA	linear
DEFINITION	HS_2020_A2_F03_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2020 Col=6 Row=K, genomic survey sequence.			
ACCESSION	AO228742			
VERSION	AO228742.1	GI:3653971		
KEYWORDS	GSS.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 473)			
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.			
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	96 (17),	9739-9744	(1999)
MEDLINE	99380589			
PubMed	10449764			
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Sequence Tagged Connector Plate: 2020 row: K column: 6 Class: BAC ends High quality sequence stop: 473.			
FEATURES	Location/Qualifiers			
source	1. 473			
	/organism="Homo sapiens"			
	/mol.type="genomic DNA"			
	/db_xref="taxon:9606"			
	/clone="Plate=2020 Col=6 Row=K"			
	/sex="male"			
	/clone.lib="CIT Approved Human Genomic Sperm Library D"			
	/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"			
BASE COUNT	157 a	101 c	88 g	126 t
ORIGIN				1 others
Query Match	84.0%;	Score 16.8;	DB 28;	Length 473;
Best Local Similarity	90.0%;	Pred. No. 4.2e+02;		
Matches	18;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
OY	1	TGTCAAAACCGTTGTGTC	20	
DB	168	TGTCAAAACCGTTGTGTC	187	
RESULT 11				
LOCUS	AI443887	508 bp	mRNA	linear
				EST 28-NOV-2001

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	COMMENT
sc44f100.y1 Gm-c1004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1004-2180 5' similar to TR:Q96502 Q96502 COL2. ;, mRNA sequence.	A1443887	A1443887				1 (bases 1 to 508)	Shoemaker, R., Kelm, P., Vodkin, L., Expanding, J., Corvett, V., Khanna, A., Bolla, B., Marr, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., C., Wyle, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ratter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	Public soybean EST Project	Unpublished
									Contact: Shoemaker R/Public Soybean EST Project

Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-553-4363 or contact via email: [cust@resgen.com](mailto:cust@resgen.com)  
Insert Length: 1480 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 393  
COLYA-NO.

FEATURES	Location/Qualifiers
source	1. .508

/organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone=GENOME SYSTEMS CLONE ID: Gm-cl004-2180"  
 /tissue\_type="root"  
 /lab\_host="X1L0-6014"  
 /clone\_11b="Gm-cl004"  
 /note=Vector: plasmidscript II XR; site\_1: EcoRI; site\_2:  
 XhoI; Root cDNA. The mRNA was isolated from entire roots  
 of 8 day old 'Williams' seedlings which were propagated on  
 paper towels with distilled water. Stratagene's cDNA  
 Synthesis Kit (catalog #200401) was used to synthesize the  
 cDNA. First-strand synthesis was performed with 5-methyl  
 dCTP, hence the ligated cDNA is hememethylated.  
 Stratagene's first-strand synthesis primer was used  
 (GAGGAGAGAGAGAGACAACTACTCTCGAG(T)-18). After  
 second-strand synthesis, the cDNA ends were 'polished'  
 with clone plu DNA polymerase, ligated to EcoRI adaptors,  
 and phosphorylated. The XhoI site within the first-strand  
 synthesis primer was restricted by digestion with XhoI;  
 all XhoI sites in the cDNA would be protected by their  
 hememethylated status. The cDNA constructs were  
 size-fractionated with a 500bp cutoff, using GIBCOBRL Life  
 Technologies' cDNA Size Fractionation column. The column  
 eluent was then ligated into Stratagene's plasmidscript II  
 XR Predigested vector (plasmidscript II SK(+)) that had been  
 digested with EcoRI and XhoI, and phosphorylated). Both  
 the white and blue colonies appear to contain recombinant  
 plasmids with cDNA inserts. Blue colonies 9n-15) have been  
 sequenced, and possess putative cDNA inserts. This library  
 was constructed by Dr. Paul Keim & Virginia H. Coryell,  
 Department of Biology, Box5640, Northern Arizona  
 University, Flagstaff, AZ 86011, Phone: 520-523-1078  
 (Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:  
 520-523-7500, email: paul.keim@nau.edu,  
 virginia.coryell@nau.edu"

BASE COUNT	95 a	193 c	118 g	102 t
ORIGIN				
Query Match		84.0%;	Score 16.8;	DB 9;
Best Local Similarity		90.0%;	Pred. No. 4.3e+02;	Length 508;
Matches 18;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps 0;				
OY	1	TGTCAAAACGGTGTGTC	20	
DB	447	TGTCAAAACGGCTGTGTC	466	

RESULT 12	CA645372	645 bp	mrna	linear	EST 23-MOV-2002
LOCUS	CA645372				
DEFINITION	CA645372				
	wreln.pK0094.c4 wreln <i>Triticum aestivum</i> cDNA clone wreln.pK0094.c4				
ACCESSION	CA645372				
VERSION	CA645372.1	GI:25223668			
KEYWORDS	EST				
SOURCE	<i>Triticum aestivum</i> (bread wheat)				
ORGANISM	<i>Triticum aestivum</i>				

REFERENCE	1 (bases 1 to 645)
AUTHORS	Tingey, S.V., Powell, W., Walters, P., Dolan, M., Hainey, C., Yuan, Z.,
TITLE	Miao, G., Caraher, N. and Hanaley, M.K.
JOURNAL	DuPont Wheat cDNA Sequence
COMMENT	Unpublished
	Contact: Scott V. Tingey

E. I. Dupont de Nemours and Company  
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA  
Tel: 302-631-2602  
Fax: 302-631-2607  
Email: Scott.V.Tingey@USA.dupont.com  
Seq primer: M13.

FEATURES	Location/Qualifiers
source	1. .645

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wreln.pk0094.c4"
/tissue_type="root"
/clone_id="wreln"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
XhoI; Wheat (Triticum aestivum L.) root; normalized from
wreln library"

```

Query Match	84.0%	Score 16.8	DB 14	Length 645
Best Local Similarity	90.0%	Pred. No. 4.5e+02		
Matches 18	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

```

QY      1 TGTCAAAAACCGTTGTGTCC 20
          ||||| | |||||
Db      584 TGTCAAAATCTGTGTGTCC 603

```

RESULT 13					
BI830009/c					
LOCUS	BI830009	669 bp	mRNA	linear	EST 04-OCT-2001
DEFINITION	60308027221 NIH_MGC_119	Homo sapiens	CDNA clone IMAGE:5171968	5'	
	mRNA sequence.				

VERSION B1830009.1 GI:15941555  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 669)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 COMMENT Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Plate: LLM11427 row: P column: 17  
 High quality sequence stop: 653.  
 Location/Qualifiers

FEATURES  
 source  
 1..669  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5171968"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH-MGC\_119"  
 /note="Organ: Brain; Vector: PCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH-MGC Library."  
 BASE COUNT 157 a 197 c 189 g 126 t  
 ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 669;  
 Best Local Similarity 90.0%; Pred. No. 4.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTC 20  
 Db 545 TGTCAAAACCGTGTGTC 526

RESULT 14  
 BU941677 875 bp mRNA linear EST 18-OCT-2002  
 LOCUS AGENCOURT\_10554416 NIH-MGC\_127 Homo sapiens CDNA clone  
 DEFINITION IMAGE:6713986 5', mRNA sequence.  
 ACCESSION BU941677  
 VERSION BU941677.1 GI:24130496  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 875)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: NCI  
 CDNA Library Preparation: Michael Brownstein Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Plate: LLM5025 row: C column: 10  
 High quality sequence stop: 302.  
 Location/Qualifiers

FEATURES

source

1..875  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6713986"  
 /tissue\_type="mixed (pool of 40 RNAs)"  
 /lab\_host="DH10B (T1-phage-resistant)"  
 /clone\_lib="NIH-MGC\_127"  
 /note="Vector: pDNR-LIB; Site\_1: SfiI (ggccatattgccc); Site\_2: SfiI (ggccgctcggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:  
 5'-AGCAGTGTATCAACGAGAGCCATTCAGCGCCGG-3' and  
 5'-ATTCTAGAGCGCGAGCGCCGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH-MGC\_126 and NIH-MGC\_128). Library created in the laboratory of T. Urdin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH-MGC Library."  
 BASE COUNT 306 a 193 c 193 g 177 t 6 others  
 ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 875;  
 Best Local Similarity 90.0%; Pred. No. 4.8e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTC 20  
 Db 839 TGTCAAAACCGTGTGTC 858

RESULT 15  
 BG335280/c 929 bp mRNA linear EST 27-FEB-2001  
 LOCUS 602403794F1 NIH-MGC\_21 Homo sapiens CDNA clone IMAGE:4541350 5', mRNA sequence.  
 DEFINITION BG335280  
 ACCESSION BG335280  
 VERSION BG335280.1 GI:13141718  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 929)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Lang Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Plate: LLM1219 row: 1 column: 23  
 High quality sequence start: 2  
 High quality sequence stop: 686.  
 Location/Qualifiers

FEATURES  
 source

1..929  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4541350"  
 /tissue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"

/clone\_11b-"NIH\_MGC\_21"  
 /note="Organ: placenta; Vector: pORB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 222 a 254 c 288 g 165 t  
 ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 929;  
 Best Local Similarity 90.0%; Pred. No. 4.8e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGTCAAAAACGCTGTGTC 20  
 ||||| ||||| |||||  
 DB 914 TGTCAAGAACCTGTGTC 895

Search completed: August 23, 2003, 12:13:20  
 Job time : 1851 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 11:11:44 ; Search time 141 Seconds  
(Without alignments)  
318.888 Million cell updates/sec

Title: US-09-927-585A-1  
Perfect score: 20  
Sequence: 1 tgcataaacgctgtctc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1517243 seqs, 1124081882 residues  
Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCY\_NEM\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCYUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEM\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_NEM\_PUB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US60\_NEM\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	12	US-10-286-387-8
2	20	100.0	20	14	US-10-292-848-18
3	20	100.0	23	12	US-10-286-387-9
4	20	100.0	23	12	US-10-286-387-10
5	20	100.0	23	12	US-10-286-387-13
6	15.8	79.0	1409	14	US-10-128-714-2427
7	15.8	79.0	1410	14	US-10-128-714-7427
8	15.8	79.0	1518	14	US-10-128-714-6427
9	15.8	79.0	1521	14	US-10-128-714-1427
10	15.8	79.0	3520	14	US-10-128-714-5427
11	15.8	79.0	3520	14	US-10-128-714-427
12	15.2	76.0	20	12	US-10-286-387-15
13	15.2	76.0	104	11	US-09-841-492-119
14	15.2	76.0	138	11	US-09-841-492-122
15	15.2	76.0	431	11	US-09-918-995-33630
16	15.2	76.0	458	12	US-10-177-390-5

17	15.2	76.0	477	9	US-09-864-761-6139	Sequence 6139, Ap
18	15.2	76.0	496	10	US-09-783-590-10792	Sequence 10792, A
19	15.2	76.0	507	13	US-10-027-632-184710	Sequence 184710,
20	15.2	76.0	525	14	US-10-255-536-71	Sequence 71, Appl
21	15.2	76.0	618	13	US-10-027-632-191550	Sequence 191550,
22	15.2	76.0	621	13	US-10-027-632-269909	Sequence 269909,
23	15.2	76.0	779	10	US-09-970-477-3	Sequence 3, Appl1
24	15.2	76.0	822	13	US-10-000-903-3	Sequence 3, Appl1
25	15.2	76.0	879	13	US-10-000-903-9	Sequence 9, Appl1
26	15.2	76.0	1116	13	US-10-000-903-5	Sequence 5, Appl1
27	15.2	76.0	1173	13	US-10-000-903-13	Sequence 13, Appl
28	15.2	76.0	1537	13	US-10-027-632-263346	Sequence 263346,
29	15.2	76.0	2368	14	US-10-133-642-1	Sequence 1, Appl1
30	15.2	76.0	2832	13	US-10-115-178-2	Sequence 2, Appl1
31	15.2	76.0	8115	10	US-09-778-516A-1	Sequence 1, Appl1
32	15.2	76.0	8115	10	US-09-778-516A-2	Sequence 2, Appl1
33	15.2	76.0	9025608	14	US-10-156-761-1	Sequence 1, Appl1
34	15	75.0	692	9	US-09-910-943-270	Sequence 270, App
35	15	75.0	758	9	US-09-910-943-358	Sequence 358, App
36	15	75.0	860	13	US-10-027-632-8841	Sequence 8841, Ap
37	14.8	74.0	338	11	US-09-803-719-1753	Sequence 1753, Ap
38	14.8	74.0	598	13	US-10-027-632-214264	Sequence 214264,
39	14.8	74.0	637	13	US-10-027-632-207791	Sequence 207791,
40	14.8	74.0	637	13	US-10-027-632-207792	Sequence 207792,
41	14.8	74.0	710	13	US-10-027-632-25868	Sequence 25868, A
42	14.8	74.0	710	13	US-10-027-632-25869	Sequence 25869, A
43	14.8	74.0	710	13	US-10-027-632-25870	Sequence 25870, A
44	14.8	74.0	726	14	US-10-081-051-75	Sequence 75, Appl
45	14.8	74.0	742	13	US-10-027-632-34048	Sequence 34048, A

## ALIGNMENTS

RESULT 1  
US-10-286-387-8  
; Sequence 8, Application US/10286387  
; Publication No. US20030143529A1  
; GENERAL INFORMATION:  
; APPLICANT: Cytic Corporation  
; TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes  
; FILE REFERENCE: gym-035CP  
; CURRENT APPLICATION NUMBER: US/10/286,387  
; CURRENT FILING DATE: 2003-02-28  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV  
US-10-286-387-8

Query Match 100.0%; Score 20; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTCAAAACCGTGTCTCC 20  
Db 1 TGTCAAAACCGTGTCTCC 20

RESULT 2  
US-10-292-848-18  
; Sequence 18, Application US/10292848  
; Publication No. US20030119042A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSADO, Pedro M.  
; APPLICANT: LOPEZ, Gemma L.  
; APPLICANT: SEITZ, Antonio M.  
; APPLICANT: ALBERDI, Maria D.  
; TITLE OF INVENTION: Method for preparing stabilized reaction mixtures, which are t

```

; TITLE OF INVENTION: partially dried, comprising at least one enzyme, reaction mixture
; FILE OF INVENTION: containing said mixtures
; FILE REFERENCE: 9256-15
; CURRENT APPLICATION NUMBER: US/10/292,848
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/ES02/00109
; PRIOR FILING DATE: 2002-3-11
; PRIOR APPLICATION NUMBER: ES P0100569
; PRIOR FILING DATE: 2001-3-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: VI-3 oligonucleotide primer
US-10-292-848-18

Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAAACCGTGTGTCC 20
Db 1 TGTCAAAAACCGTGTGTCC 20

RESULT 3
US-10-286-387-9
; Sequence 9, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
; APPLICANT: Cytoc Corporation
; TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
; FILE REFERENCE: cym-035CP
; CURRENT APPLICATION NUMBER: US/10/286,387
; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-9

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAAACCGTGTGTCC 20
Db 1 TGTCAAAAACCGTGTGTCC 20

RESULT 4
US-10-286-387-10
; Sequence 10, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
; APPLICANT: Cytoc Corporation
; TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
; FILE REFERENCE: cym-035CP
; CURRENT APPLICATION NUMBER: US/10/286,387
; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial

; FEATURE:
; OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-10

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAAACCGTGTGTCC 20
Db 1 TGTCAAAAACCGTGTGTCC 20

RESULT 5
US-10-286-387-13
; Sequence 13, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
; APPLICANT: Cytoc Corporation
; TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
; FILE REFERENCE: cym-035CP
; CURRENT APPLICATION NUMBER: US/10/286,387
; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-13

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCAAAAACCGTGTGTCC 20
Db 1 TGTCAAAAACCGTGTGTCC 20

RESULT 6
US-10-128-714-2427/c
; Sequence 2427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Weng
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskhin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2427
; LENGTH: 1409
; TYPE: DNA
```



ORGANISM: Aspergillus fumigatus  
US-10-128-714-2427

Query Match 79.0%: Score 15.8; DB 14; Length 1409;  
Best Local Similarity 89.5%: Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GTCAAAACCGTTGTGTC 20  
Db 769 GCCAAACCGTTGTGTC 751

RESULT 7  
US-10-128-714-7427/c  
Sequence 7427, Application US/10128714  
Publication No. US20030119013A1  
GENERAL INFORMATION:

APPLICANT: Jiang, Bo  
APPLICANT: Hu, Wengqi  
APPLICANT: Tishkoff, Daniel  
APPLICANT: Zamudio, Carlos  
APPLICANT: Eroszhkin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
PRIORITY FILING DATE: 2002-04-23  
PRIORITY FILING DATE: 2001-04-23  
PRIORITY FILING DATE: 2001-04-23  
PRIORITY FILING DATE: 2001-04-27  
PRIORITY FILING DATE: 2001-04-27  
PRIORITY FILING DATE: 2001-06-05  
PRIORITY FILING DATE: 2001-06-05  
PRIORITY FILING DATE: 2001-07-09  
PRIORITY FILING DATE: 2001-07-09  
PRIORITY FILING DATE: 2001-08-31  
PRIORITY FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7427  
LENGTH: 1410  
TYPE: DNA  
ORGANISM: Aspergillus fumigatus  
US-10-128-714-7427

Query Match 79.0%: Score 15.8; DB 14; Length 1410;  
Best Local Similarity 89.5%: Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GTCAAAACCGTTGTGTC 20  
Db 770 GCCAAACCGTTGTGTC 752

RESULT 8  
US-10-128-714-6427/c  
Sequence 6427, Application US/10128714  
Publication No. US20030119013A1  
GENERAL INFORMATION:

APPLICANT: Jiang, Bo  
APPLICANT: Hu, Wengqi  
APPLICANT: Tishkoff, Daniel  
APPLICANT: Zamudio, Carlos  
APPLICANT: Eroszhkin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
PRIORITY FILING DATE: 2002-04-23  
PRIORITY FILING DATE: 2001-04-23  
PRIORITY FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: US 60/287,066  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/295,890  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/303,899  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 60/316,362  
PRIOR FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6427  
LENGTH: 1518  
TYPE: DNA  
ORGANISM: Aspergillus fumigatus  
US-10-128-714-6427

Query Match 79.0%: Score 15.8; DB 14; Length 1518;  
Best Local Similarity 89.5%: Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GTCAAAACCGTTGTGTC 20  
Db 878 GCCAAACCGTTGTGTC 860

RESULT 9

US-10-128-714-1427/c  
Sequence 1427, Application US/10128714  
Publication No. US20030119013A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Bo  
APPLICANT: Hu, Wengqi  
APPLICANT: Tishkoff, Daniel  
APPLICANT: Zamudio, Carlos  
APPLICANT: Eroszhkin, Alexey M  
APPLICANT: Lemieux, Sebastien M  
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
FILE REFERENCE: 10182-018-999  
CURRENT APPLICATION NUMBER: US/10/128,714  
PRIORITY FILING DATE: 2002-04-23  
PRIORITY FILING DATE: 2001-04-23  
PRIORITY FILING DATE: 2001-04-23  
PRIORITY FILING DATE: 2001-04-27  
PRIORITY FILING DATE: 2001-04-27  
PRIORITY FILING DATE: 2001-06-05  
PRIORITY FILING DATE: 2001-06-05  
PRIORITY FILING DATE: 2001-07-09  
PRIORITY FILING DATE: 2001-07-09  
PRIORITY FILING DATE: 2001-08-31  
PRIORITY FILING DATE: 2001-08-31  
NUMBER OF SEQ ID NOS: 8603  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1427  
LENGTH: 1521  
TYPE: DNA  
ORGANISM: Aspergillus fumigatus  
US-10-128-714-1427

Query Match 79.0%: Score 15.8; DB 14; Length 1521;  
Best Local Similarity 89.5%: Pred. No. 2.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GTCAAAACCGTTGTGTC 20  
Db 878 GCCAAACCGTTGTGTC 860

RESULT 10  
US-10-128-714-5427/c  
Sequence 5427, Application US/10128714  
Publication No. US20030119013A1  
GENERAL INFORMATION:

```

; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5427
; LENGTH: 3518
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-5427

```

```

Query Match      79.0%; Score 15.8; DB 14; Length 3518;
Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      2 GTCAAAACCGTGTGTC 20
      1 ||||| |||||
Db      1878 GCCAAAACCTGTGTCC 1860

```

```

RESULT 11
US-10-128-714-427/c
; Sequence 427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 427
; LENGTH: 3520
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-427

```

Query Match 79.0%; Score 15.8; DB 14; Length 3520;

```

Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      2 GTCAAAACCGTGTGTC 20
      1 ||||| |||||
Db      1877 GCCAAAACCTGTGTCC 1859

```

```

RESULT 12
US-10-286-387-15
; Sequence 15, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
; APPLICANT: Cytic Corporation
; TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
; FILE REFERENCE: cym-035CP
; CURRENT APPLICATION NUMBER: US/10/286,387
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-15

```

```

Query Match      76.0%; Score 15.2; DB 12; Length 20;
Best Local Similarity 85.0%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1 TGTCAAAACCGTGTGTC 20
      1 ||||| |||||
Db      1 TGTCAAAACCGTGTGTC 20

```

```

RESULT 13
US-09-941-492-119/c
; Sequence 119, Application US/09941492
; Publication No. US20030027250A1
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd
; APPLICANT: Garcia-Blanco, Mariano M.
; APPLICANT: Puttaraju, Madalah
; APPLICANT: Mansfield, Gary S.
; TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN
; FILE REFERENCE: A31304-BAE (072874.0156)
; CURRENT APPLICATION NUMBER: US/09/941,492
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/838,858
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/756,096
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; LENGTH: 104
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Binding domain of human papilloma virus P17M
US-09-941-492-119

```

Query Match 76.0%; Score 15.2; DB 11; Length 104;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTCAAAACCGTGTGTCC 20  
||||||| || |||||  
Db 25 TGTCAAAAGCCACTGTGTCC 6

## RESULT 14

US-09-941-492-122/c  
; Sequence 122, Application US/09941492  
; Publication No. US20030027250A1  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, Lloyd  
; APPLICANT: Garcia-Blanco, Mariano M.  
; APPLICANT: Puttaraju, Madalah  
; APPLICANT: Mansfield, Gary S.  
; TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN  
; TITLE OF INVENTION: SPLICESOME MEDIATED RNA TRANS-SPLICING  
; FILE REFERENCE: A31304-BAE (072874.0156)  
; CURRENT APPLICATION NUMBER: US/09/941,492  
; CURRENT FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: 09/838,858  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 09/756,096  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: 09/158,863  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 09/133,717  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: 09/087,233  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 08/766,354  
; PRIOR FILING DATE: 1996-12-13  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122  
; LENGTH: 138  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Binding domain of human papilloma virus PTM  
US-09-941-492-122

Query Match 76.0%; Score 15.2; DB 11; Length 138;  
Best Local Similarity 85.0%; Pred. No. 3.1e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTCAAAACCGTGTGTCC 20  
||||||| || |||||  
Db 60 TGTCAAAAGCCACTGTGTCC 41

## RESULT 15

US-09-918-995-33830/c  
; Sequence 33830, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 33830  
; LENGTH: 431  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-918-995-33830

Query Match 76.0%; Score 15.2; DB 11; Length 431;  
Best Local Similarity 85.0%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTCAAAACCGTGTGTCC 20  
||||||| || |||||  
Db 297 TGTCAAAAGAGTGTGTCC 278

Search completed: August 23, 2003, 13:53:16  
Job time : 147 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:18:48 ; Search time 50 seconds  
(without alignments)  
176.553 Million cell updates/sec

Title: US-09-927-585A-1

Perfect score: 20  
Sequence: 1 tgcataaacctgtgtcc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/lna/PCUTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/lna/backfills1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	15.2	76.0	29	1	US-08-495-743-27 Sequence 27, Appl
C 2	15.2	76.0	29	1	US-08-495-739-27 Sequence 27, Appl
C 3	15.2	76.0	29	1	US-08-495-741-27 Sequence 27, Appl
C 4	15.2	76.0	29	3	US-08-062-023-27 Sequence 27, Appl
5	15.2	76.0	456	3	US-09-056-105-29 Sequence 29, Appl
6	15.2	76.0	456	4	US-09-430-201-6 Sequence 6, Appl
7	15.2	76.0	519	3	US-08-860-165-11 Sequence 11, Appl
8	15.2	76.0	519	4	US-09-359-382-11 Sequence 11, Appl
9	15.2	76.0	525	4	US-09-669-751-71 Sequence 71, Appl
10	15.2	76.0	570	5	PCT-US94-05085A-1 Sequence 1, Appl
11	15.2	76.0	570	5	PCT-US94-05085-1 Sequence 1, Appl
12	15.2	76.0	776	1	US-08-216-233C-1 Sequence 1, Appl
13	15.2	76.0	779	4	US-09-210-168-3 Sequence 3, Appl
14	15.2	76.0	790	1	US-08-117-083-7 Sequence 9, Appl
15	15.2	76.0	801	3	US-08-860-165-9 Sequence 9, Appl
16	15.2	76.0	801	4	US-09-359-382-9 Sequence 9, Appl
17	15.2	76.0	822	4	US-09-485-885-3 Sequence 3, Appl
18	15.2	76.0	879	4	US-09-485-885-9 Sequence 9, Appl
19	15.2	76.0	1000	1	US-07-965-274-3 Sequence 1, Appl
20	15.2	76.0	1007	1	US-08-663-871-1 Sequence 1, Appl
21	15.2	76.0	1107	4	US-08-664-894-1 Sequence 1, Appl
22	15.2	76.0	1116	4	US-09-485-885-5 Sequence 5, Appl
23	15.2	76.0	1173	4	US-09-485-885-13 Sequence 13, Appl
24	15.2	76.0	2368	2	US-08-520-933-1 Sequence 1, Appl
25	15.2	76.0	2368	4	US-09-285-040-1 Sequence 1, Appl
C 26	15.2	76.0	2368	6	5262177-1 Patent No. 5262177
C 27	15.2	76.0	2550	1	US-08-188-228-53 Sequence 53, Appl

C 28	15.2	76.0	2550	1	US-08-332-643-47 Sequence 47, Appl
C 29	15.2	76.0	2550	1	US-08-332-638-53 Sequence 53, Appl
C 30	15.2	76.0	2832	3	US-08-819-177-2 Sequence 2, Appl
C 31	15.2	76.0	3136	1	US-08-188-228-41 Sequence 41, Appl
C 32	15.2	76.0	3136	1	US-08-332-638-41 Sequence 41, Appl
33	15.2	76.0	7833	1	US-08-074-879-9 Sequence 9, Appl
34	15.2	76.0	7833	1	US-08-468-057A-9 Sequence 1, Appl
35	15.2	76.0	7904	1	US-08-316-239B-1 Sequence 1, Appl
36	15.2	76.0	7904	1	US-08-316-239B-2 Sequence 2, Appl
37	15.2	76.0	7904	1	US-08-410-005-1 Sequence 1, Appl
38	15.2	76.0	7904	3	US-08-829-140-1 Sequence 1, Appl
39	15.2	76.0	7904	3	US-09-560-579A-1 Sequence 1, Appl
40	15.2	76.0	10803	3	US-09-080-044-1 Sequence 1, Appl
41	15.2	76.0	10803	4	US-09-531-857A-1 Sequence 1, Appl
42	14.8	74.0	1028	3	US-08-990-823-69 Sequence 69, Appl
43	14.8	74.0	1028	4	US-09-477-135A-69 Sequence 69, Appl
44	14.8	74.0	1107	4	US-09-000-094-19 Sequence 19, Appl
45	14.8	74.0	1128	4	US-09-000-094-21 Sequence 21, Appl

## ALIGNMENTS

RESULT 1  
US-08-495-743-27/C  
Sequence 27, Application US/08495743  
Patent No. 5654416  
GENERAL INFORMATION:  
APPLICANT: Thomas J. Cummins, Susan M. Alwood  
APPLICANT: Lynn Berngayer, John B. Findlay  
APPLICANT: John W.H. Sutherland, Joanne H. Kerschner  
TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,  
METHODS AND TEST KITS FOR  
TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO  
TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eastman Kodak Company, Patent Legal Staff  
STREET: 343 State Street  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14650 - 2201  
MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)  
COMPUTER READABLE FORM:  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: PC-8 (Word for Windows)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/495,743  
FILING DATE: 08-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/062,023  
FILING DATE: 14-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Tucker, J. Lanny  
REGISTRATION NUMBER: 27,678  
REFERENCE/DOCKET NUMBER: 67271A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 722-9332  
TELEFAX: (716) 477-4646  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 nucleotides  
TYPE: Nucleic acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Primer for HPV DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Synthetically prepared

IMMEDIATE SOURCE: Same  
PUBLICATION INFORMATION: No. 5654416  
US-08-495-743-27

Query Match 76.0%; Score 15.2; DB 1; Length 29;  
Best Local Similarity 85.0%; Pred. No. 36;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTCC 20  
||||| |||  
DB 20 TGTCAAAACCGACGTGTCC 1

RESULT 2  
US-08-495-739-27/c  
Sequence 27, Application US/08495739  
Patent No. 5702901

## GENERAL INFORMATION:

APPLICANT: Thomas J. Cummins, Susan M. Atwood  
APPLICANT: Lynn Bergmeyer, John B. Findlay  
APPLICANT: John W.H. Sutherland, Joanne H. Kerschner  
TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,  
TITLE OF INVENTION: METHODS AND TEST KITS FOR  
TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO  
TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS  
TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES  
NUMBER OF SEQUENCES: 65

## CORRESPONDENCE ADDRESS:

ADDRESS: Eastman Kodak Company, Patent Legal Staff  
STREET: 343 State Street  
CITY: Rochester  
STATE: New York

COUNTRY: U.S.A.  
ZIP: 14650 - 2201

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS Version 3.3

SOFTWARE: PC-8 (word for windows)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/495,739

FILING DATE: 08-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/062,023

FILING DATE: 14 MAY 1993

ATTORNEY/AGENT INFORMATION:

NAME: Tucker, J. Lanny

REGISTRATION NUMBER: 27,678

REFERENCE/DOCKET NUMBER: 67271A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 722-9332

TELEFAX: (716) 477-4646

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 nucleotides

TYPE: Nucleic acid

STRANDEDNESS: Single

MOLECULE TYPE: Primer for HPV DNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE: Synthetically prepared

IMMEDIATE SOURCE: Same

PUBLICATION INFORMATION: No. 5702901e

US-08-495-739-27

Query Match 76.0%; Score 15.2; DB 1; Length 29;  
Best Local Similarity 85.0%; Pred. No. 36;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TGTCAAAACCGTGTGTCC 20  
||||| |||

DB 20 TGTCAAAACCGACGTGTCC 1

RESULT 3  
US-08-495-741-27/c  
Sequence 27, Application US/08495741  
Patent No. 5733751

## GENERAL INFORMATION:

APPLICANT: Thomas J. Cummins, Susan M. Atwood

APPLICANT: Lynn Bergmeyer, John B. Findlay

APPLICANT: John W.H. Sutherland, Joanne H. Kerschner

TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,

TITLE OF INVENTION: METHODS AND TEST KITS FOR

TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO

TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS

TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES

NUMBER OF SEQUENCES: 65

## CORRESPONDENCE ADDRESS:

ADDRESS: Eastman Kodak Company, Patent Legal Staff

STREET: 343 State Street

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14650 - 2201

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS Version 3.3

SOFTWARE: PC-8 (word for windows)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/495,741

FILING DATE: 08-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/062,023

FILING DATE: 19930514

ATTORNEY/AGENT INFORMATION:

NAME: Tucker, J. Lanny

REGISTRATION NUMBER: 27,678

REFERENCE/DOCKET NUMBER: 67271A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 722-9332

TELEFAX: (716) 477-4646

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 nucleotides

TYPE: Nucleic acid

STRANDEDNESS: Single

MOLECULE TYPE: Primer for HPV DNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE: Synthetically prepared

IMMEDIATE SOURCE: Same

PUBLICATION INFORMATION: No. 5733751e

US-08-495-741-27

Query Match 76.0%; Score 15.2; DB 1; Length 29;  
Best Local Similarity 85.0%; Pred. No. 36;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TGTCAAAACCGTGTGTCC 20  
||||| |||  
DB 20 TGTCAAAACCGACGTGTCC 1

RESULT 4  
US-08-062-023-27/c  
Sequence 27, Application US/08062023  
Patent No. 6174668  
GENERAL INFORMATION:  
APPLICANT: Thomas J. Cummins, Susan M. Atwood

APPLICANT: Lynn Bergmeyer, John B. Findlay  
APPLICANT: John W.H. Sutherland, Joanne H. Kerschner  
TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,  
TITLE OF INVENTION: METHODS AND TEST KITS FOR  
TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO  
TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eastman Kodak Company, Patent Legal Staff  
STREET: 343 State Street  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14650 - 2201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: PC-8 (Word for Windows)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/062,023  
FILING DATE: 19930514  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Tucker, J. Lanay  
REGISTRATION NUMBER: 27,678  
REFERENCE/DOCKET NUMBER: 67271A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 722-9332  
TELEFAX: (716) 477-4646  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 nucleotides  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Primer for hpv DNA  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE: Synthetically prepared  
IMMEDIATE SOURCE: Same  
PUBLICATION INFORMATION: No. 6174668e  
US-08-062-023-27

Query Match 76.0%; Score 15.2; DB 3; Length 29;  
Best Local Similarity 85.0%; Pred. No. 36;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGTCAAAACCGTGTGTCC 20  
||||| ||| |||  
Db 20 TGTCAAAACCGACTGTGTCC 1

RESULT 5  
US-09-056-105-29  
Sequence 29, Application US/09056105  
Patent No. 6287569  
GENERAL INFORMATION:  
APPLICANT: KIPPS, THOMAS J.  
APPLICANT: MU, YUNOI  
TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR  
TITLE OF INVENTION: PROCESSING  
FILE REFERENCE: 233/221  
CURRENT APPLICATION NUMBER: US/09/056,105  
CURRENT FILING DATE: 1998-04-06  
EARLIER APPLICATION NUMBER: 60/043,467  
EARLIER FILING DATE: 1997-04-10  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 29  
LENGTH: 456  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-056-105-29

Query Match 76.0%; Score 15.2; DB 3; Length 456;  
Best Local Similarity 85.0%; Pred. No. 51;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGTCAAAACCGTGTGTCC 20  
||||| ||| |||  
Db 316 TGTCAAAACCGACTGTGTCC 335

RESULT 6  
US-09-430-201-6  
Sequence 6, Application US/09430201  
Patent No. 631373  
GENERAL INFORMATION:  
APPLICANT: Eckert, Richard L.  
APPLICANT: Crish, James F.  
TITLE OF INVENTION: Tissue Specific Promoters and Transgenic Animals for  
FILE REFERENCE: CASE-04022  
CURRENT APPLICATION NUMBER: US/09/430,201  
CURRENT FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/106,495  
PRIOR FILING DATE: 1998-10-30  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 456  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-430-201-6

Query Match 76.0%; Score 15.2; DB 4; Length 456;  
Best Local Similarity 85.0%; Pred. No. 51;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGTCAAAACCGTGTGTCC 20  
||||| ||| |||  
Db 316 TGTCAAAACCGACTGTGTCC 335

RESULT 7  
US-08-860-165-11  
Sequence 11, Application US/08860165A  
Patent No. 6004557  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, Stirling John  
APPLICANT: COX, John Cooper  
APPLICANT: WEBB, Elizabeth Ann  
APPLICANT: FRAZER, Ian  
TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS  
FILE REFERENCE: 17227/130  
CURRENT APPLICATION NUMBER: US/08/860,165A  
CURRENT FILING DATE: 1997-09-22  
EARLIER APPLICATION NUMBER: PCT/AU95/00868  
EARLIER FILING DATE: 1995-12-20  
EARLIER APPLICATION NUMBER: AU PN0157  
EARLIER FILING DATE: 1994-12-20  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(516)  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion  
US-08-860-165-11

Query Match 76.0%; Score 15.2; DB 3; Length 519;  
Best Local Similarity 85.0%; Pred. No. 52;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTCC 20  
|||||  
Db 151 TGTCAAAAGCCACGTGTCC 170

## RESULT 8

US-09-359-382-11  
Sequence 11, Application US/09359382  
Patent No. 6306397  
GENERAL INFORMATION:  
APPLICANT: EDWARDS, Stirling John  
APPLICANT: COX, John Cooper  
APPLICANT: WEBB, Elizabeth Ann  
APPLICANT: FRAZER, Ian  
TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS  
FILE REFERENCE: 017227/0148  
CURRENT APPLICATION NUMBER: US/09/359,382  
CURRENT FILING DATE: 1999-07-23  
EARLIER APPLICATION NUMBER: US 08/860,165  
EARLIER FILING DATE: 1997-09-22  
EARLIER APPLICATION NUMBER: PCT/AU95/00868  
EARLIER FILING DATE: 1995-12-20  
EARLIER APPLICATION NUMBER: AU PN0157/94  
EARLIER FILING DATE: 1994-12-20  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Human papillomavirus type 16  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(516)  
US-09-359-382-11

Query Match 76.0%; Score 15.2; DB 4; Length 519;  
Best Local Similarity 85.0%; Pred. No. 52;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTCC 20  
|||||  
Db 151 TGTCAAAAGCCACGTGTCC 170

## RESULT 9

US-09-669-751-71  
Sequence 71, Application US/09669751  
Patent No. 6551575  
GENERAL INFORMATION:  
APPLICANT: Greenspan, Ralph J.  
TITLE OF INVENTION: Methods for Identifying Compounds for  
TITLE OF INVENTION: Motion Sickness, Vertigo and other Disorders Related to  
TITLE OF INVENTION: Balance and the Perception of Gravity  
FILE REFERENCE: P-NI 3864  
CURRENT APPLICATION NUMBER: US/09/669,751  
CURRENT FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US 60/168,579  
PRIOR FILING DATE: 1999-12-02  
NUMBER OF SEQ ID NOS: 261  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 71  
LENGTH: 525  
TYPE: DNA  
ORGANISM: Drosophila  
US-09-669-751-71

Query Match 76.0%; Score 15.2; DB 4; Length 525;  
Best Local Similarity 85.0%; Pred. No. 52;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTCC 20  
|||||  
Db 244 TGTCAAAACCGTGTGTCC 263

## RESULT 10

PCT-US94-05085A-1  
Sequence 1, Application PC/TUS9405085A  
GENERAL INFORMATION:  
APPLICANT: Janice T. Brown  
TITLE OF INVENTION: HUMAN PAPILLOMAVIRUS DETECTION ASSAY  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Baxter Diagnostics Inc.  
STREET: One Baxter Parkway, Building DP-3E  
CITY: Deerfield  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Macintosh System 7.0  
SOFTWARE: Macintosh Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05085A  
FILING DATE: N/A  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/058,920  
FILING DATE: May 6, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark Buonaiuto  
REGISTRATION NUMBER: 31,593  
REFERENCE/DOCKET NUMBER: BA-4448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708/948-2537  
TELEFAX: 708/948-2642  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
HYPOTHETICAL: no  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Papovaviridae, Human papilloma virus  
STRAIN: 16  
FEATURE:  
NAME/KEY: Portion of viral genome coding for E6/E7 polypeptides.  
PUBLICATION INFORMATION:  
AUTHORS: Seedorf, K., Kramerer, G., Durst, M.,  
AUTHORS: Suhai, S., and Rowekamp, W.  
TITLE: Human Papillomavirus Type 16 DNA Sequence  
JOURNAL: Virology  
VOLUME: 145  
ISSUE:  
PAGES: 181-185  
DATE: 1985  
PCT-US94-05085A-1

Query Match 76.0%; Score 15.2; DB 5; Length 570;  
Best Local Similarity 85.0%; Pred. No. 53;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGTCAAAACCGTGTGTCC 20  
|||||  
Db 338 TGTCAAAAGCCACGTGTCC 357



RESULT 11  
PCT-US94-05085-1  
Sequence 1, Application PC/TUS9405085  
GENERAL INFORMATION:  
APPLICANT: Janice T. Brown  
TITLE OF INVENTION: HUMAN PAPILLOMAVIRUS DETECTION ASSAY  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Baxter Diagnostics Inc.  
STREET: One Baxter Parkway, Building DP-3E  
CITY: Deerfield  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Macintosh System 7.0  
SOFTWARE: Macintosh Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05085  
FILING DATE: 06-MAY-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/058,920  
FILING DATE: May 6, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark Buonaiuto  
REGISTRATION NUMBER: 31,593  
REFERENCE/DOCKET NUMBER: BA-4448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708/948-2537  
TELEFAX: 708/948-2642  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
HYPOTHETICAL: no  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Papovaviridae, Human papilloma virus  
SPRAIN: 16  
FEATURE:  
NAME/KEY: Portion of viral genome coding for E6/E7 polypeptides.  
PUBLICATION INFORMATION:  
AUTHORS: Seedorf, K., Krammer, G., Durst, M.,  
AUTHORS: Subal, S., and Rowekamp, W.  
TITLE: Human Papillomavirus Type 16 DNA Sequence  
JOURNAL: Virology  
VOLUME: 145  
ISSUE:  
PAGES: 181-185  
DATE: 1985  
PCT-US94-05085-1  
Query Match 76.0%; Score 15.2; DB 5; Length 570;  
Best Local Similarity 85.0%; Pred. No. 53;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TGTCAAAACCGTGTGTC 20  
DB 338 TGTCAAAAGCCACTGTGTC 357

RESULT 12  
US-08-216-233C-1  
Sequence 1, Application US/08216233C  
Patent No. 5506105  
GENERAL INFORMATION:

APPLICANT: HAYDOCK, PAUL V.  
TITLE OF INVENTION: IN SITU ASSAY OF AMPLIFIED INTRACELLULAR  
TITLE OF INVENTION: RNA TARGETS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DADE INTERNATIONAL INC.  
STREET: 1717 DEERFIELD ROAD  
CITY: DEERFIELD  
STATE: ILLINOIS  
COUNTRY: USA  
ZIP: 60015  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/216,233C  
FILING DATE: 22-MAR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/808,456  
FILING DATE: 10-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: TYMESON, CYNTHIA G.  
REGISTRATION NUMBER: 34,745  
REFERENCE/DOCKET NUMBER: BA-4203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (305) 222-6423  
TELEFAX: (305) 222-6686  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 776 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-216-233C-1  
Query Match 76.0%; Score 15.2; DB 1; Length 776;  
Best Local Similarity 85.0%; Pred. No. 55;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TGTCAAAACCGTGTGTC 20  
DB 337 TGTCAAAAGCCACTGTGTC 356

RESULT 13  
US-09-210-168-3  
Sequence 3, Application US/09210168  
Patent No. 635424  
GENERAL INFORMATION:  
APPLICANT: Lorincz, Attila T.  
TITLE OF INVENTION: ASSESSMENT OF HUMAN PAPILLOMA VIRUS-RELATED DISEASE  
FILE REFERENCE: 26294005052  
CURRENT APPLICATION NUMBER: US/09/210,168  
CURRENT FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: U.S. 60/082,167  
PRIOR FILING DATE: 1998-04-17  
PRIOR APPLICATION NUMBER: U.S. 60/070,486  
PRIOR FILING DATE: 1998-01-05  
PRIOR APPLICATION NUMBER: U.S. 60/069,426  
PRIOR FILING DATE: 1997-12-12  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 3  
LENGTH: 779  
TYPE: DNA  
ORGANISM: Human papillomavirus  
FEATURE:  
OTHER INFORMATION: E6/E7-HPV16  
US-09-210-168-3

Query Match 76.0%; Score 15.2; DB 4; Length 779;  
Best Local Similarity 85.0%; Pred. No. 55;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 TGTCAAAACCGTGTGTCC 20  
||||| ||| |||  
Db 344 TGTCAAAAGCCACGTGTGCC 363

RESULT 14  
US-08-117-083-7  
; Sequence 7, Application US/08117083  
; Patent No. 5719054  
; GENERAL INFORMATION:  
; APPLICANT: Boursnell, Michael E.  
; APPLICANT: Ingils, Stephen C.  
; APPLICANT: Munro, Alan J.  
; TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human  
; TITLE OF INVENTION: Papilloma Virus Proteins  
; NUMBER OF SEQUENCES: 70  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Walter H. Dreger  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/117,083  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Walter H.  
; REGISTRATION NUMBER: 24,190  
; REFERENCE/DOCKET NUMBER: A-58783  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; FAX: 910 277299  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 790 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; US-08-117-083-7

Query Match 76.0%; Score 15.2; DB 1; Length 790;  
Best Local Similarity 85.0%; Pred. No. 55;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTCAAAACCGTGTGTCC 20  
||||| ||| |||  
Db 342 TGTCAAAAGCCACGTGTGCC 361

RESULT 15  
US-08-860-165-9  
; Sequence 9, Application US/08860165A  
; Patent No. 6004557  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, Stirling John  
; APPLICANT: COX, John Cooper  
; APPLICANT: WEBB, Elizabeth Ann  
; APPLICANT: FRAZER, Ian  
; TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS

FILE REFERENCE: 17227/130  
; CURRENT APPLICATION NUMBER: US/08/860,165A  
; CURRENT FILING DATE: 1997-09-22  
; EARLIER APPLICATION NUMBER: PCT/AU95/00868  
; EARLIER FILING DATE: 1995-12-20  
; EARLIER APPLICATION NUMBER: AU PNO157  
; EARLIER FILING DATE: 1994-12-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 801  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(798)  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion  
; US-08-860-165-9

Query Match 76.0%; Score 15.2; DB 3; Length 801;  
Best Local Similarity 85.0%; Pred. No. 55;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 TGTCAAAACCGTGTGTCC 20  
||||| ||| |||  
Db 337 TGTCAAAAGCCACGTGTGCC 356

Search completed: August 23, 2003, 12:15:08  
Job time : 51 secs